

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 12, 2005, 04:00:43 ; Search time 5145 Seconds  
(without alignments)  
10397.387 Million cell updates/sec

Title: US-10-642-531-1  
Perfect score: 1104  
Sequence: 1 atgcgcgtcgaagaagcctag.....cgcagagccttcgcatga 1104

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenBank:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the local score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	878.4	79.6	1132	8	AY059083 Arabidops
2	878.4	79.6	1441	8	AF039406 Arabidops
3	878.4	79.6	1483	8	AY035017 Arabidops
4	873.6	79.1	1457	6	AR269261 Arabidops
5	873.6	79.1	1457	8	ATW3312 Arabidops
6	639.2	57.9	1537	8	AJ630634 Arabidops
7	545.8	49.4	1500	8	AY026039 Oryza sat
8	542.8	49.1	1422	8	AK069157 Oryza sat
9	524.8	47.5	1424	8	BT009355 Triticum
10	524.6	47.5	1602	8	AF038586 Zea mays
11	517.4	46.3	1380	8	AF038585 Zea mays
12	511.4	46.3	1355	8	AK100033 Oryza sat
13	442	40.0	1405	8	AF323611 Oryza sat
14	324.2	29.4	899	6	AR566929 Sequence
15	324	29.3	86022	8	ATAC011623 Arabidops
16	208.2	18.9	371	6	BD267332 Composite
17	208.2	18.9	371	6	AR566843 Sequence
18	191.8	17.4	110000	8	CR382129_28 Continuation (29 o
19	191.4	17.3	94704	8	AC020580 Arabidops

C 20	177.8	16.1	129306	8	NCB24N11	BX284751 Neurospor
C 21	177.8	16.1	199388	8	NCB14D6	AL561173 Neurospor
C 22	171.2	15.5	3326	8	AY589088	AY589088 Glycine m
C 23	163.8	14.8	110000	2	AP006500_05	Continuation (6 of
C 24	154.4	14.0	3415	8	AY589089	AY589089 Glycine m
C 25	137	12.4	109399	8	AP003749	AP003749 Oryza sat
C 26	137	12.4	119039	2	AP004383	AP004383 Oryza sat
C 27	135	12.2	111484	8	AC144475	AC144475 Medicago
C 28	135	12.2	111486	8	AC136140	AC136140 Medicago
C 29	135	12.2	111486	8	AC147536	AC147536 Medicago
C 30	135	12.2	135914	8	AC146757	AC146757 Medicago
C 31	130.8	11.8	94647	8	AC145372	AC145372 Medicago
C 32	130.8	11.8	94656	8	AC138579	AC138579 Medicago
C 33	121.4	11.0	236	6	AX645560	AX645560 Sequence
C 34	121.4	11.0	236	6	AX676721	AX676721 Sequence
C 35	120.2	10.9	624	6	AR549025	AR549025 Sequence
C 36	120.2	10.9	1536	8	AX489451	AX489451 Sequence
C 37	118.2	10.7	158714	8	AC082644	AC082644 Oryza sat
C 38	104.2	9.4	33733	8	SPAC644	AL355012 S.pombe c
C 39	104.2	9.4	110000	8	CR382135_11	Continuation (12 o
C 40	99.4	9.0	1149	10	RNU27456	U27456 Rattus norv
C 41	99.4	9.0	1654	10	AP043070	AP043070 Mus muscu
C 42	99.4	9.0	1822	10	BC046595	BC046595 Mus muscu
C 43	90.2	8.2	3212	5	BC070978	BC070978 Xenopus 1
C 44	89.2	8.1	102197	8	CR382126_25	Continuation (26 o
C 45	86.8	7.9	110000	8	AE016818_11	Continuation (12 o

## ALIGNMENTS

RESULT 1	AY059083	1132 bp	mRNA	linear	PLN 18-SEP-2002
LOCUS	AY059083				
DEFINITION	Arabidopsis thaliana putative pyruvate dehydrogenase kinase				
ACCESSION	(AY059083) mRNA, complete cds.				
VERSION	AY059083.1	GI:16323389			
KEYWORDS	FLI CDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	1 (bases 1 to 1132)				
AUTHORS	Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Egu, P., Lee, J.M., Toriumi, M., Yu, G., Brooks, S., Chao, Q., Chen, H., Karlin-Neumann, G., Kim, C., Lam, B., Miranda, M., Nguyen, M., Palm, C.J., Shin, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.				
TITLE	Arabidopsis Open Reading Frame (ORF) Clones				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1132)				
AUTHORS	Yamada, K., Banh, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shin, P., Southwick, A., Tracy, S.E., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-OCT-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA				
COMMENT	The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RFLV cDNAs (RFLV cDNA : RIKEN Arabidopsis Full-length cDNA): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.				

The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the pLNU (ORF) clones using the RFLV cDNAs: Yamada, K., Banh, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M.,





The Saik, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Liu, S. X., Sakano, H., Pham, P. K., Bann, J., Chung, M. K., Goldsmith, A. D., Lee, J. M., Quach, H. L., Tang, C. C., Toriumi, M., Yu, G., Bower, L., Chen, H., Cherk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M. C., Miranda, M., Nguyen, M., Palm, C. J., Shinn, P., Southwick, A., Davis, R. W., Ecker, J. R. and Theologis, A.

Yamada, K. (SSP/PGSC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGSC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

# FEATURES

source  
1.1483  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/clone="RAFL05-09-E15 (R10102)"  
/ecotype="Columbia"  
/note="This clone is in a modified pBluescript vector (lambda Zap) as a XhoI/SerI insert."  
gene  
1.1483  
/gene="At3g06483"  
5'UTR  
1.143  
/gene="At3g06483"  
144.1244  
/gene="At3g06483"  
CDS  
144.1244  
/gene="At3g06483"  
/codon\_start=1  
/evidence="experimental"  
/product="putative pyruvate dehydrogenase kinase"  
/protein\_id="AAK59522.1"  
/db\_xref="GI:14334688"  
/translation="MAVKKACEMFPKSLIEDVHKCMKQTVSLRYMFEFSKPTER NLISAOFLHKEIPKIVARPAIELOTPGLSDPAVLKRVLYLESFDDMAFPIK DSGDEKQFMKIPKIVARPAIELOTPGLSDPAVLKRVLYLESFDDMAFPIK IRLMIGQVELHNPDPPLTVGYIHTKMSMEVARASDARSICREVGSAPEINITY GDSPTFPPVYELHNPDPPLTVGYIHTKMSMEVARASDARSICREVGSAPEINITY SDEGGIARSGLRIFTLYLSTARNLEEDVDIGIADVVTNAGYRGIPISRLYARY FGGLDIIISMEGSDPAVYHLSRLSDSDEPLP"  
1245.1483  
/gene="At3g06483"  
1467  
/gene="At3g06483"  
/note="artifact within poly A tail"

## ORIGIN

Query Match 79.6%; Score 878.4; DB 8; Length 1483;  
Best Local Similarity 87.9%; Pired. No. 1.4e-248;  
Matches 970; Conservative 0; Mismatches 131; Indels 3; Gaps 1;  
Db 1 ATGGCGGTGAAGAGGCTACGAGATGTTTGAAGAGCTGATGAGAGCGTTTCAACA 60  
144 ATGGCAGTGAAGAGGCTGCGAAATGTTCCGAGAGATTGATGAGAGATTCAACA 203  
QY 61 TGGGATGATGAGAGAGCGGCGCTGAGCCTCAGGTACATGATGAGAGTTCCGTTCA 120  
Db 204 TGGGGTTGATGAGAGAGCGGCTGTTACCTTAAGTATCATATGAGAGTTGGTTCCAA 263  
QY 121 CCACATGAGAGAACTTCTGATCTCGGCGAGTTTCTTCAACAGAGCTTCCGATTGG 180  
Db 264 CTTACTGAGAGAAATTTTGTGATTTCTGCTCAGTTTGTGACATAGAGAGCTTCCGATTGG 323  
QY 181 ATCGGAGGCGTGCATCGAATCGAGAGCTGCTTATGAGCTCTGAGAGAACTCGCC 240  
Db 324 GTGCGCAGAGAGCGATCGAATCGAGAGCTTCTTATGAGCTCTGATTAACCTGCG 383  
QY 241 GTCTTGAAGATGAAGATTGATGATGAGAGTCAATTCAGGAGCATGAGAGCGTTTCTGAG 300  
Db 384 GTTTGAAGGTCGGGATTGTTGATTTGGAATCTTCAAGGACATGAGAGCATTTCTGAG 443

QY 301 ATCAAGATATCTGCTGATGAGAGAGATTCAACAGATGATCAAGGCTTTAAAGTAAG 360  
Db 444 ATTAAGATTCGGGTACAGAGAGAGATTTCATCTCATATATTAAGCTTCAAGTAAG 503  
QY 361 CAAACAGAGTGTGTTCCATGATGAGTGTGAGTGTGAGAACCACTGAGAGAGTAAGAA 420  
Db 504 CATACAAATGTGTGTTCCATGATGAGTGTGAGTGTGAGAACCACTGAGAGAGTAAGAA 562  
QY 421 CTCTAGAGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
Db 563 --TTCTGAGAAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 620  
QY 481 GGGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
Db 621 GGGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 680  
QY 541 ACAGTGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
Db 681 ACAGTGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 740  
QY 601 GATGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
Db 741 GATGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 800  
QY 661 GATCCAGTTTCACTTTCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
Db 801 GATCCAGTTTCACTTTCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 860  
QY 721 GTCAAGAACTCTCTCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
Db 861 GTCAAGAACTCTCTCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 920  
QY 781 CCAATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
Db 921 CCAATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 980  
QY 841 GGTGAGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
Db 981 GGTGAGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1040  
QY 901 AGAAACCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
Db 1041 AGAAACCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1100  
QY 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
Db 1101 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1160  
QY 1021 CAGATCATATCATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
Db 1161 CAGATCATATCATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1220  
QY 1081 GATTCGAGAGAGCTTTGCCATGA 1104  
Db 1221 GATTCGAGAGAGCTTTGCCATGA 1244

RESULT 4  
AR269261  
LOCUS AR269261  
DEFINITION Sequence 1 from patent US 6500670.  
ACCESSION AR269261  
VERSION AR269261.1  
KEYWORDS GI:29700268  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1457)  
AUTHORS Zou, J. and Taylor, D. C.  
TITLE Plant pyruvate dehydrogenase kinase gene  
JOURNAL Patent: US 6500670-A 1-31-DEC-2002;  
FEATURES Location/Qualifiers



```

61 TGGGATGATGATGAGCAGACGGGCGTGAAGCTCAGATGATCATGATGAGATTCGGTCACT 120
164 TGGGATGATGATGAGCAGACGGGCGTGAAGCTCAGATGATCATGATGAGATTCGGTCACT 223
121 CCCACTGAGAGAAACCTTCGTATCTCGGCGAGATTCCTTCAAGAGACCTCCGATTCGG 180
224 CCACTGAGAGAAACCTTCGTATCTCGGCGAGATTCCTTCAAGAGACCTCCGATTCGG 283
181 ATCGGAGAGCGTGCATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTCC 240
284 GTGCGAGAGAGCGATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTCC 343
241 GTCTGAGAGATGAGATTCGATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTCC 300
344 GTTTGAGAGATGAGATTCGATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTCC 403
301 ATCAAGATGATGATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTCC 360
404 ATTAAGATGATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTCC 463
361 CACAACACGATGATGATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTCC 420
464 CATACATGATGATGATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTCC 522
421 CTGACGAGAGATGATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTCC 480
523 --TTTGGAAATCTGATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTCC 580
481 GGGATCGGATGATGATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTCC 540
581 GGGATCGGATGATGATGAGATTCGAGACCGTCTGAGATTCGATTAATCCAAATCCACGCTTCAT 640
541 ACAATGAGATGATGATGATGAGATTCGATTAATCCAAATCCACGCTTCAT 600
641 ACAATGAGATGATGATGAGATTCGATTAATCCAAATCCACGCTTCAT 700
601 GATGAGATGATGATGATGAGATTCGATTAATCCAAATCCACGCTTCAT 660
701 GATGAGATGATGATGATGAGATTCGATTAATCCAAATCCACGCTTCAT 760
661 GATGAGATGATGATGATGAGATTCGATTAATCCAAATCCACGCTTCAT 720
761 GATGAGATGATGATGATGAGATTCGATTAATCCAAATCCACGCTTCAT 820
721 GTCAGAGATGATGATGATGAGATTCGATTAATCCAAATCCACGCTTCAT 780
821 GTCAGAGATGATGATGATGAGATTCGATTAATCCAAATCCACGCTTCAT 880
781 CCAATCGGATGATGATGAGATTCGATTAATCCAAATCCACGCTTCAT 840
881 CCAATCGGATGATGATGAGATTCGATTAATCCAAATCCACGCTTCAT 940
841 GGTGAGATGATGATGAGATTCGATTAATCCAAATCCACGCTTCAT 900
941 GGTGAGATGATGAGATTCGATTAATCCAAATCCACGCTTCAT 1000
901 AGAAACCCCTGAGAGATTCGATTAATCCAAATCCACGCTTCAT 960
1001 AGAAACCCCTGAGAGATTCGATTAATCCAAATCCACGCTTCAT 1060
961 GGTGAGATGATGATGAGATTCGATTAATCCAAATCCACGCTTCAT 1020
1061 GGTGAGATGATGATGAGATTCGATTAATCCAAATCCACGCTTCAT 1120
1021 CAGATCATATCAGATGAGATTCGATTAATCCAAATCCACGCTTCAT 1080
1121 CAGATCATATCAGATGAGATTCGATTAATCCAAATCCACGCTTCAT 1180
1081 GACTCGAGAGATTCGATTAATCCAAATCCACGCTTCAT 1104
1181 GATTGCAAGAGATTCGATTAATCCAAATCCACGCTTCAT 1204

```

```

RESULT 6
Aj630654
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
1 Duplico, B., Martin, I. and Labrador, E.
A pyruvate dehydrogenase kinase is expressed in Cicer arietinum
etiolated epicotyls
Unpublished
2 (bases 1 to 1537)
Labrador, E.
Direct Submission
Submitted (10-MAR-2004) Labrador E., Dpto. Fisiologia Vegetal,
Univ. Salamanca, Campus Miguel de Unamuno, E-37007, SPAIN
Location/Qualifiers
1.1537
/organism="Cicer arietinum"
/mol_type="mRNA"
/cultivar="Pedrosillano"
/db_xref="taxon:3827"
/clone="CampDX"
/tissue_type="etiolated epicotyls"
/clone_id="CAN-5"
/dev_stage="5 days old seedlings"
1..1537
/gene="Pdk"
147..1250
/gene="Pdk"
/EC_number="2.7.1.99"
/codon_start=1
/product="pyruvate dehydrogenase kinase"
/db_xref="GI:45720178"
/db_xref="GI:45720178"
/translation="MAAIEAFSKSLIEVHKMGCLKOTGVSLRYMERGSKPTDNL
ISAOFLKELAIRAIRAIELETLPIYGLSGSDAIVKRVVDSFRDRIKSPKNTL
DEKETEVIKAIKVRHNVVPTMALGVQQLKKGAKRNWEDPEIHQFLRPLSRI
GIRMLIGHVEHLNPNPPYVGIHTMSVEVARNASVDARSICUREYGSADPINT
YGDPEPTPEVPAHLHLMVFEVLVNSLAAVDERVDSKXSPRIIIVADLEBVTIK
ISDEGGIARSGLPKIFTYLYSTRNPLDEHEDIGVDSVTTMAGYGLPISRLYAR
YFGDGLQIISMEGYTDAYLHLSRLGDSQEPDP"

ORIGIN
Query Match 57.9%; Score 639.2; DB 8; Length 1537;
Best Local Similarity 74.2%; Pred. No. 1.le-177;
Matches 825; Conservative 0; Mismatches 278; Indels 9; Gaps 1;

2 TGGCGTGAAGAAGCTGACGAGATGTTTGAAGAGCTTGATGAGAGAGCTTACAGAT 61
139 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 198
62 GGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121
199 GGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 258
122 CCACTGAGAGAAACCTTCGTATCTCGGCGAGATTCCTTCAAGAGACCTCCGATTCGG 181
259 CTACTGAGAGAAATTTGATTTCTGACAGATTCCTTCAAGAGAACTTGATGAGAG 318
182 TCGGAGAGCGTGCATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTCCG 241
319 TTGCTGAGAGAGCTATGAGCTTGAACCTTCCTGATGATGATGATGATGATGATGAT 378
242 TCTTGAAGGTAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 301

```



Db 379 TATTGAAGTTAGGATTTGGTATGTGATTCATTCGCCGATATTCGATCTTTCCCTGAAA 438  
 302 TCAAGGATCTGCTGATGAGAAAGAGTTCAACAGATGATCAAGGCGCTTAAAGTAAAGC 361  
 Db 439 TCAAGATATTAATGATGAAAAAGAGTTCACTGAGTTATCAAGGCGCATCAAGTGAAC 498  
 Qy 362 ACAACAACGTTGTTCCCATGATGAGCTCTGGGTGGAACCAAGCTGAGAAAGAAATGAAC 421  
 Db 499 ACAACATGTGTATCTCAACATGCGCTTAGGTGTGACAGTTGAAAAAGGCTTGAAAC 558  
 Qy 422 -----TCAAGAAAGCTTGATGATGATTCATCAGTTCTTGTATCGCTTCTACTGT 472  
 Db 559 CAATATGTTATGAAAGATTTGTTGAGATTCATCAGTTCTTGTATCGCTTCTACTGT 618  
 Qy 473 CTGCTATGAGGATCCGATGCTTATGCGGAGATGATGATGATGATGATGATGATGATGAT 532  
 Db 619 CAAGAAATGGAATTCGATGCTTATGAGGAGATGATGATGATGATGATGATGATGATGAT 678  
 Qy 533 CACTTCAACAGTGGGTTACATACACACCAAGATGTCCTTATGAGAGTGGCAAGATG 592  
 Db 679 CACCCATGTTGGGTTATATATATCAATAAATGTCCTGTCGAGTGGCGAAGATG 738  
 Qy 593 CTATGTAAGATGCAAGGCTGATTTGTTCAAGATATGTTCTGCTCCGAGATTAACA 652  
 Db 739 CCAGTGAAGATGCAAGGCTGATTTGTTCAAGATATGTTCTGCTCCGAGATTAACA 798  
 Qy 653 TATATGCGATCCAGTTTCACTTCCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 712  
 Db 799 TTTATGATGCTTCACTTCTTCCGATGTCGATGTCGATGTCGATGTCGATGTCGATGAT 858  
 Qy 713 ATGATGATGCAAGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAT 772  
 Db 859 TTGAGTTGTTAAGATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAT 918  
 Qy 773 TTGACCAACCAATCCGATGATGTTGCTGATGATGATGATGATGATGATGATGATGAT 832  
 Db 919 TTTCACTCCCATTAAGATATGTTGCTGATGATGATGATGATGATGATGATGATGAT 978  
 Qy 833 CAGATGAAGTGAAGTATACGGAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAT 892  
 Db 979 CAGATGAAGTGAAGTATACGGAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGAT 1038  
 Qy 893 GCACCTCAAGAAACCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGAT 952  
 Db 1039 GATCTCCCAAGAAACCACTTGAAGATGATGATGATGATGATGATGATGATGATGAT 1098  
 Qy 953 CTATGCTGTTATGTTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAT 1012  
 Db 1099 CTATGCTGTTATGTTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAT 1158  
 Qy 1013 GAGATTTGAGATCATATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1072  
 Db 1159 GATGATTTCAAAATATCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1218  
 Qy 1073 GTCTTGAAGCTGCAAGAGCTTTGCTGATGA 1104  
 Db 1219 GTTTGGAGATTCACAAAGACCTTTACCTGA 1250

RESULT 7  
 AY026039 1500 bp mRNA linear PLN 14-FEB-2001  
 LOCUS Oryza sativa subsp. indica pyruvate dehydrogenase kinase 1 mRNA,  
 DEFINITION complete cds.  
 ACCESSION AY026039  
 VERSION AY026039.1 GI:12829951  
 KEYWORDS Oryza sativa (indica cultivar-group)  
 SOURCE Oryza sativa (indica cultivar-group)  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE 1 (bases 1 to 1500)

AUTHORS Yau, C.P., Zhuang, C.X., Yip, W.K. and Zee, S.Y.  
 TITLE Molecular characterization of a cDNA encoding putative pyruvate  
 dehydrogenase kinase in rice  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1500)  
 AUTHORS Yau, C.P., Zhuang, C.X., Yip, W.K. and Zee, S.Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-DEC-2000) Department of Botany, The University of  
 Hong Kong, Room 76/06, Kadoorie Biological Science Building, Hong  
 Kong, Hong Kong

FEATURES  
 source  
 1. 1500  
 /organism="Oryza sativa (indica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="IR36"  
 /sub\_species="indica"  
 /db\_xref="taxon:39946"  
 /dev\_stage="first and second meiosis"  
 55..1146  
 /note="OS-PDK1"  
 /codon\_start=1  
 /product="pyruvate dehydrogenase kinase 1"  
 /protein\_id="AAK01947.1"  
 /db\_xref="GI:12829952"  
 /translation="MASEVBARAVAEVGRWQKQGVSLRYMEFGSPETERNLL  
 SAQFLQKELPIRLARALESLPFLSKPALKVLDVDFDIIFYEPEVRNND  
 ELAPQIMIKIVRHNHVVPTMALGVQOLKNEVYRKIPTAFDEIHEFDEYMSRI  
 GIRMUGQVLAHDPDPEPGVGLITNELSPQVNOAASBPASICLRGASPEIDI  
 YQDPTTFPYSVSHLHMLPELVKNSLRVQERKNSDQVPRVITVADGADVTIK  
 VSDGGGIRPSGLPRIFTYLYSTAKNPMDPSESEVYAGYGLPISLVARYFEG  
 DLIISMEGYGDVAYLHLSRLDPSERP"

CDS  
 55..1146  
 /note="OS-PDK1"  
 /codon\_start=1  
 /product="pyruvate dehydrogenase kinase 1"  
 /protein\_id="AAK01947.1"  
 /db\_xref="GI:12829952"  
 /translation="MASEVBARAVAEVGRWQKQGVSLRYMEFGSPETERNLL  
 SAQFLQKELPIRLARALESLPFLSKPALKVLDVDFDIIFYEPEVRNND  
 ELAPQIMIKIVRHNHVVPTMALGVQOLKNEVYRKIPTAFDEIHEFDEYMSRI  
 GIRMUGQVLAHDPDPEPGVGLITNELSPQVNOAASBPASICLRGASPEIDI  
 YQDPTTFPYSVSHLHMLPELVKNSLRVQERKNSDQVPRVITVADGADVTIK  
 VSDGGGIRPSGLPRIFTYLYSTAKNPMDPSESEVYAGYGLPISLVARYFEG  
 DLIISMEGYGDVAYLHLSRLDPSERP"

ORIGIN  
 Query Match 49.4%; Score 545.8; DB 8; Length 1500;  
 Best Local Similarity 69.8%; Pred. No. 5.1e-150;  
 Matches 777; Conservative 0; Mismatches 312; Indels 24; Gaps 2;

Qy 4 GCAGTGAAGAGCTAGCGAGATGTTTCAAGAGCTTATCGAGACGTTCAAGATG 63  
 Db 46 GGGGGGAGATGCGTCGAGACCGGTCGCGAGGCGCGTGGCGAGAGAGTGGAGTGG 105  
 Qy 64 GATGATGAAGAGAGCGGCGGCTGAGCTCAAGTATCATATGAGTGGTTCACCTCC 123  
 Db 106 GGGAGCATGAAGAGAGCGGCGGCTGAGCTCAAGTATCATATGAGTGGTTCACCTCC 165  
 Qy 124 ACTGAGAGAACTTGTGATCTGCGGCACTTCTTCAAGAGAGCTTCCGATTCGATC 183  
 Db 166 ACGAGACGGAACCTGTGCTCTCCGCGAGTTCCTGCAAGAGAGCTCCCAATCCGATC 225  
 Qy 184 GCGAGGCGTGCATTCGAGACGTCGAGAGCTTATGAGCTTCTGAGAAACCTGCGCTC 243  
 Db 226 GCGCGCGCGCGCTCGAGCTCGAGTCTCTCCCTTGGGCTCTCCCGCAAGCCGCGATC 285  
 Qy 244 TTGAAGTAAAGATTTGATGATGAGTCAATTCAGGAGATGAGAGCTTCTGAGATC 303  
 Db 286 CTCGAAGTCGCGGATTTGATGATGAGTCTTTCGCGAGATGAGATGATGATGATGATG 345  
 Qy 304 AAGGATATGCTGATGAGAGAGTTCACACAGATGATGAGAGCTTTAAAGAGCAC 363  
 Db 346 AAGGATATGAGAGAGAGTTCGCTTTTACGAGATGATGAGAGCTTTAAAGAGCAC 405  
 Qy 364 AACCAAGTATTCCTGATGATGAGTCTGAGTGAACCAAGTGAAGAAAG----- 412  
 Db 406 AATATATGTTGTTCAACATGAGTCAAGAGTTCACACAGTCAAGAGTATCCG 465  
 Qy 413 -GAATGAACCTTACGAAAAAGCTTGAATGATTCATCACTTCTTATCTTACTTG 471  
 Db 466 ACAAGAGAGATCCCACTGATTTGAGAGATTCACAGAGTTCCTTACAGATTTCTACATG 525  
 Qy 472 TCTCGATATGAGATCCGATGCTTATGCGGCGAGAGTGTGATGATGATGATGATGATGAT 531  
 Db 526 TCAAGGATTTGAGATCCGATGCTTATGAGGAGAGATGAGCTTGTGATGATGATGATGATG 585

Oy	532	CCACTTCACACAGTGGTTACATACACCAAGATGTCCTATGAGGTGGCAAGAT	591
Db	586	GAGCTGGGCTGATAGGCTCTATATACAGAAATTTCCCTATACAGGTGGCTCAAGCT	645
Oy	552	GCTAGTGAAGATGCAAGTGCATTTGTTTCAGAGATATGTTCTGCTCCGAGATTAAC	651
Db	646	GCCATGTGAAGATGCCGCTTATTTGTTGAGGGAATATGATCAAGCTCCGAGATTTAC	705
Oy	652	ATAATGCGATTCACAAAGTTTCACTTTCCGATGTTCCGACCCATTTGCATCTATGTC	711
Db	706	ATCTATGAGACCCCACTTTTACATTTTCATATGTTTATATACATCTATCTATGCTA	765
Oy	712	TATGAGTATGTCAGAACTCTCTCCGTCTGTCCAGAGCCGTTTGTGACTGTATAGG	771
Db	766	TTTGATTTGTGAGAACTCTCTGTGCTGATGAGAGAGATATATGAAATTCGATTA	825
Oy	772	GTTGACACCAATCCGATATGTTGCTGATGATGATGATGATGATGATGATGATGATG	831
Db	826	GATGTTCTCCAGTTGAAATTAATGCTGATGAGGAGAGATGATGATGATGATGATG	885
Oy	832	TCAGATGAGGTGAGATATACCGAAGCGGTCTCCCTAAATATTCATTAATCTTAC	891
Db	886	AGTATGAGAGTGTGAGATTCAGAGAGCGGCTTCCAGAAATTTTCAATATCTATAT	945
Oy	892	AGCACTGACAGAAACCACTTGAAGATGATGATGATGATGATGATGATGATGATG	951
Db	946	AGCACTGACAGAAATTCACCT-----GATATGATGATGATGATGATGATGATG	993
Oy	952	ACTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1011
Db	994	ACTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1053
Oy	1012	GGAGATTTGACATATATCATGAGAGATGACGAGCTGATGATGATGATGATGATG	1071
Db	1054	GATGACCTGCAAAATCATCTCATATGAGAGATGAGAGATGATGATGATGATGATG	1113
Oy	1072	CGTTTGAGAGCTGCGAGAGCTTTCCTTCAATGA	1104
Db	1114	CGTTTGAGAGATTCGAGAGACCTTTCCTTGA	1146
RESULT 8			
AK069157			
LOCUS			
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:U023007C01, full		
ACCESSION	AK069157		
VERSION	AK069157.1		
KEYWORDS	FLI CDNA; CAP trapper.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.		
REFERENCE			
AUTHORS	1 The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otsu, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nishikawa, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Mura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN; Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arikawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Komano, H., Miyazaki, A., Otsu, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.		
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from		

# JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

japonica rice  
Science 301 (5631), 376-379 (2003)  
22752273  
12869764  
2 (bases 1 to 1422)  
Adachi, J., Aizawa, K., Arikawa, T., Arikawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Komano, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Mura, J., Narikawa, R., Namiki, T., Narikawa, R., Nishikawa, R., Nishikawa, R., Nomura, K., Numasaki, R., Ohneda, E., Ohta, Y., Ohtsuki, K., Oka, M., Ooka, H., Otsu, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Suzuki, K., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.

## TITLE JOURNAL

Direct Submission  
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@ias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)  
This clone is one of the 28k full-length cDNA clones from japonica rice.  
URL: <http://cdna01.dna.affrc.go.jp/cDNA/>  
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and Yamamoto, M.

## COMMENT

FAIS Genome Sequencing & Analysis Group; Otsu, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Mura, J., Mizuno, K., Narikawa, R., Nishikawa, R., Nishikawa, R., Nomura, K., Sugiyama, A., Suzuki, K., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K.  
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Komano, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishikawa, K., Nomura, K., Numasaki, R., Ohta, Y., Ohtsuki, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A., and Hayashizaki, Y.

## FEATURES source

1. 1422  
Location/Qualifiers  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultiyar="Nidpobare"  
/db\_xref="taxon:39947"  
/clone="U023007C01"

## ORIGIN

Query Match 49.1%; Score 542.2; DB 8; Length 1422;  
Best Local Similarity 70.1%; Pred. No. 5,9e-149;  
Matches 768; Conservative 0; Mismatches 303; Indels 24; Gaps 2;



QY 22 GAGATGTTTTCGAGAGCTTGATCGAGAGCTTCACAGATGGGATGATGAGAGCAGACG 81  
 DB 6 GAGCCGGTGGCGAGGGCGGTGGCGAGAGGTGGGGAGGTGGGGAGCATGAGAGCAGACG 65  
 QY 82 GGGGTGAGCTTCAGGTACATGATGAGAGTTGGTTCCATCTCCATCTGAGAGAACTTCTG 141  
 DB 66 GGGGTGAGCTTCAGGTACATGATGAGAGTTGGTTGGTGGCGCCCGAGGAGAGCTGCTG 125  
 QY 142 ATCTGGCGGAGTTTCTTCGAGAGAGCTTCGAGTTGAGTTCGAGGCGGTGGAGTGA 201  
 DB 126 CTCTCCGCGAGTTCTTCGAGAGAGAGCTTCGAGTTGAGTTCGAGGCGGTGGAGTGA 185  
 QY 202 CTCGAGAGCTTCGAGTTGAGTTCGAGAGAGCTTCGAGTTGAGTTCGAGGAGTGA 261  
 DB 186 CTCGAGAGCTTCGAGTTGAGTTCGAGAGAGCTTCGAGTTGAGTTCGAGGAGTGA 245  
 QY 262 TATGTGAGTTCATTCAGGAGCATGAGAGCTTCGAGTTGAGTTCGAGGAGTGA 321  
 DB 246 TACTGTGAGTTCATTCAGGAGCATGAGAGCTTCGAGTTGAGTTCGAGGAGTGA 305  
 QY 322 AAGAGTTCAGAGATGATGATGAGAGCTTCGAGTTGAGTTCGAGGAGTGA 381  
 DB 306 CTCGTTTATTCAGAGATGATGATGAGAGCTTCGAGTTGAGTTCGAGGAGTGA 365  
 QY 382 ATGGCTCTGGGTGAGAGCTGAGAGAG-----GAATGAACTCTAGAG 429  
 DB 366 ATGGAGCTGGAGTTTCAAGAGCTCAAGAGAGAGAGTTCGAGTTGAGTTCGAGGAGTGA 425  
 QY 430 AAGCTTGAAGATTCATCAGTTTCTTATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 489  
 DB 426 GCATTTGAGAGATTCAGAGAGTTTCTTATGAGATTCATCAGTTTCTTCTTCTTCTTCTT 485  
 QY 490 ATGGCTTATGGGAGAGATGATGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 549  
 DB 486 ATGGCTTATGGGAGAGATGATGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 545  
 QY 550 TACATACAGAGAGATGATGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 609  
 DB 546 CTCATTAATACAGAGATGATGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 605  
 QY 610 TCGATTTGTTTTCAGAGATGATGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 669  
 DB 606 TCTATTTGTTTTCAGAGATGATGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 665  
 QY 670 TTTCACTTTCCGATGATGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 729  
 DB 666 TTTCACTTTCCGATGATGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 725  
 QY 730 TCTCTCCGCTGATGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 789  
 DB 726 TCTCTCCGCTGATGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 785  
 QY 790 ATCATTTGTTTTCAGAGATGATGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 849  
 DB 786 ATTATCTGTTGATGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 845  
 QY 850 ATACGAGAGAGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 909  
 DB 846 ATACGAGAGAGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 905  
 QY 910 CTTCGAGAGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 969  
 DB 906 CTT-----GATATGATGCTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 953  
 QY 970 TATGTCTGCTTATGATGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 1029  
 DB 954 TATGTCTGCTTATGATGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 1013  
 QY 1030 TCCATGAGAGATGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 1089  
 DB 1014 TCTATGAGAGATGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGTTGAGTTCGAGGAGTGA 1073  
 QY 1090 GAGCTTTTTCGATGA 1104

Db 1074 GAGCCCTTGCCTTGA 1088  
 RESULT 9  
 BT009355 1424 bp mRNA 1line PLN 20-JUN-2003  
 LOCUS Trifolium aestivum clone wlm96.pk0020.d2:fls, full insert mRNA  
 DEFINITION  
 sequence.  
 BT009355  
 ACCESSION  
 BT009355  
 VERSION  
 BT009355.1 GI:32128906  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Trifolium aestivum (bread wheat)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticeae; Triticum.  
 1 (bases 1 to 1424)  
 Jolley, S.V., Wolters, P., Powell, W., Dolan, M., Miao, G.-H.,  
 Caraher, N.R., Hanafey, M.K. and Hainey, C.F.  
 Direct Submission  
 Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and  
 Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,  
 USA  
 FEATURES  
 source  
 1..1424  
 /organism="Trifolium aestivum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4565"  
 /clone="wlm96.pk0020.d2:fls"  
 ORIGIN  
 Query Match 47.5%; Score 524.8; DB 8; Length 1424;  
 Best Local Similarity 68.8%; Pred. No. 8; 5e-144;  
 Matches 762; Conservative 0; Mismatches 322; Indels 24; Gaps 2;  
 QY 7 GTGAGAGAGCTAGCAGAGATGTTTTCGAGAGAGCTTATGAGAGCTTCAGAGAGCTTCAGAGTGGGA 66  
 DB 122 GGGATGATGCGCGTGGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 181  
 QY 67 TCGATTAAGCAGAGCGGAGCTGAGCTCAGCTTACATGATGAGAGTTCGAGTTTCTTCTTCTTCTTCTT 126  
 DB 182 AGCATTAACAGAGCGGAGCTGAGCTCAGCTTACATGATGAGAGTTCGAGTTTCTTCTTCTTCTTCTT 241  
 QY 127 GAGAGAACTTCTGATCTCGGAGAGTTCCTTCAAGAGAGCTTCGAGTTGAGTTCGAGTGGAGCGG 186  
 DB 242 GACCGAGAGCTGCTCTTCTGAGAGAGTTCCTTCAAGAGAGCTTCGAGTTTCTTCTTCTTCTTCTT 301  
 QY 187 AGGCTGCGATGAGAGCTGAGAGAGCTTATGAGCTTCTGAGAGAGCTTCGAGTTTCTTCTTCTTCTT 246  
 DB 302 CCGCGGCGCTGAGAGCTGAGAGAGCTTCTTCTGAGAGAGCTTCGAGTTTCTTCTTCTTCTTCTT 361  
 QY 247 AAGTGAAGATGATGATGAGTTCAGAGTTCAGAGAGCTTCGAGTTTCTTCTTCTTCTTCTTCTT 306  
 DB 362 AAGTGAAGATGATGATGAGTTCAGAGTTCAGAGAGCTTCGAGTTTCTTCTTCTTCTTCTTCTT 421  
 QY 307 GATACGCTGAGAGAGAGATTCAGAGATGATGAGTTCAGAGAGCTTCGAGTTTCTTCTTCTTCTTCTT 366  
 DB 422 AATCGGAGAGAGAGCTTCGAGTTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 481  
 QY 367 AAGTGTTCCTGATGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 414  
 DB 482 AAGTGTTCCTGATGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 541  
 QY 415 ATGAACTCTAGAGAGAGCTTCGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTC 474  
 DB 542 AAGAGAGAGAGAGAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTC 601  
 QY 475 CGATAGAGAGAGAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTC 534  
 DB 602 AAGAT 661  
 QY 535 CTTCACAGAGTGGTTACATACACCAAGATGCTCTTATGAGAGTGGCAAGAGATGCT 594

```

Db      662  CCTGCGCTCATGAGCCCTCATCAATAAATGTGCCCCATACAGGTAGCTCAATTCCT
Qy      595  AGTGAAGATGCAAGTGCATTTGTTTCAGAGATGTGGTTGCTCCGAGATAACATA
Db      722  AGTGAAGATGCGCGTTTCCATTTGTATGAGGAAATACGATACGCTCCGACATCAACTT
Qy      655  TATGCGCATCCCAAGTTTCCATTTTCCGATGTTCCGACCCATTTGCATCTTATGAT
Db      782  TATGAGATCAAAATTTAACTTTCCATATGTCATGTCATCTTCACTCATCTGTT
Qy      715  GAGTTAGTCAAGAACTCTCTCCGTCGTCCCAAGACGGTTTGTGACTGTATAGGAT
Db      842  GAACGTGGGAAAATTTCCCTTCGTGACAGAGAAATATATGAAATTTGATTAAGAT
Qy      775  GCACCAACCAATCCGATCATTTGTTGCTGATGGAATCCAAAGTGTATCAATAAGTCTCA
Db      902  GTTCTCTCTGTCAGAAATTTATGTTGCTGATGGAACAGAGATGTTACTATTAAGTCACT
Qy      835  GATGAAGGTGAGGTATATACCGAAGAGCGTCTCCCTAAATAATTCACCTTACCTAGAGC
Db      962  GATGAAGCGGTGAAATACGGAAGATGGGCTCCCTGAAATTTCACTATCTTTACAGC
Qy      895  ACTGCAAGAAACCACTTTGAAGAAGATGTGACTTGGAAACCGTGAATGTTCCCTGACT
Db      1022  ACTCAAAAGATCTAATCTT-----GATATGAGGCGCTTACTGGAAGAGTACT
Qy      955  ATGCTGCTTATGTTATGCTGCTCTATGATGCTTGTATGCTGCTTATTTGGTGA
Db      1070  ATGCTGCTGATATGTTTGTGACTTCAAGTATGCTGCTTATGCGCAATTTGGTGGC
Qy      1015  GATTGCAATCATATCATGGAAGATACGGAAGTATGCTTACTTCACTTGTCTGT
Db      1130  GACCTACAAATCATATCATGAGAGATACGCGCACGACCTTACCTCCACTGTACGA
Qy      1075  CTGAGAGACTCGAGAGCGCTTGCCAT 1102
Db      1190  CTGGAGACTCTGGAGAGCCATTTGCTT 1217

RESULT 10
LOCUS   AF038586                1602 bp    mRNA    linear    PLN 15-OCT-1998
DEFINITION  Zea mays pyruvate dehydrogenase kinase isoform 2 mRNA, nuclear gene
ACCESSION AF038586
VERSION   AF038586.1 GI:3695004
KEYWORDS
SOURCE    Zea mays
ORGANISM  Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1602)
AUTHORS   Thelen,J.J., Muzynski,M.G., Miernyk,J.A. and Randall,D.D.
TITLE      Molecular analysis of two pyruvate dehydrogenase kinases from maize
JOURNAL   J. Biol. Chem. 273 (41), 26518-26623 (1998)
MEDLINE   98434573
PUBMED    9756901
2 (bases 1 to 1602)
REFERENCE Thelen,J.J., Miernyk,J.A. and Randall,D.D.
AUTHORS   Direct Submission
TITLE      Submitted (15-DEC-1997) Biochemistry, University of Missouri, 117
JOURNAL   Schweitzer Hall, Columbia, MO 65211, USA
FEATURES
         source
         1..1602
         /organism="Zea mays"
         /mol_type="mRNA"
         /db_xref="taxon:4577"
         78..1172
         /EC_number="2.7.1.99"
         /function="catalyzes the phosphorylation of pyruvate
         dehydrogenase"
CDS

```

```

/note="PDK2"
/codon_start=1
/product="pyruvate dehydrogenase kinase isoform 2"
/protein_id="AAC63962.1"
/db_xref="GI:3695005"
/transtrans="MASEPBARVAEVEVRMSMKTQVTLRYHMEFGRPTORULL
SAGFLKEPPIPARRALDLSIPFLSKPAILKRVWDYLDSPDIRFPEVSRND
ELAFITQIMWVNRNNVPTMALGVQQLKELGSRKVPFEFSDIDEFLFYSRI
GIRMLIGQVALDPEPQVIGLINTRLSPIQVAAACEDASVCLAEYGAPDINI
YDPPNTPFPYVTLHLHMLFELVKISLRAVQRVYNSDQVDPVRIIVADGSEDTIK
VDEGGGIPRSGLPPIFTLYLSTAKNPELDLPNTERVYMAVYGGPLISRLYAVYFG
GDLQIISMEGYDAVILHSLRGDSERP"
ORIGIN
Query Match 47.5%; Score 524.6; DB 8; Length 1602;
Best Local Similarity 68.8%; Pred. No. 9,9e-144;
Matches 760; Conservative 0; Mismatches 324; Indels 21; Gaps 2;
Qy      12  GAAAGCTAGCGAGATGTTTTCGAGAAGCTGTATCGAGAGCTTCAAGATGGGATGCT
Db      77  GATGGCGTGGAGCCGCTGGCGCGGCGGTGGCGAGAGGTGGCGCGCTGGGAGAGCT
Qy      72  GAAGCAGAGCGGCGTGGAGCGCTGAGTACATGATGAGTTCGCTTCCACTCCAGTGAAG
Db      137  GAAGCAGAGCGGCGGTGACCTCGGTATCATGATGAGTTCGCTTCCGCGCCAGCCAGC
Qy      132  AAACCTTGATCTGGCGCACTTCTTCAAGAAGCTTCCGATTGGAATGCGAGCG
Db      197  CAACCTGCTCTCCGCGCACTTCTTCAAGAAGCTTCCGATCCGCTTCCGCGCG
Qy      192  TCGATCGAATCGAGAGCGCTGCTTATGCGCTCTGAGAAGACCTGCGCTTGAAGT
Db      257  CGCGCTGAGCTCGAGCTGCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT
Qy      252  AAGATGTTGATGTTGAGTCAATTCAGGAGATGAGAGCTTCTGATCAAGATAC
Db      317  GCGGAGCTGTTACTTGAATCATTCAGGAGATGAGATCTTCTGAAAGTGAAGAGCG
Qy      312  TGCTGATGAGAAGGTTCAACAGATGATGATGATGATGATGATGATGATGATGATGAT
Db      377  GAAAGAGAGCTGCTTTCAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      372  GGTTCCTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
Db      437  GGTTCCTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
Qy      432  GCTT-----GATGATTCATCACTTCTTCTGATGCTTCTTCTGATGCTTCTGAT
Db      497  GGTTCATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      480  AGGATCCGATGCTTATCGGAGAGATGATGATGATGATGATGATGATGATGATGATGAT
Db      557  TGGATCCGATGCTTATCGGAGAGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      540  CACAGTGGTATCATACACCAAGATGCTTCTTATGAGAGTGGCAAGATGCTAGTGA
Db      617  TGTATAGGCTCATCAATACAGATATATCCCAATCCAGTGGCTCAAGTGGTGTGTA
Qy      600  AGATGCAAGTGCATTTGTTTCAAGAGTATGTTGCTTCCGAGATTAACATATATG
Db      677  AGATGCCGCTTCTGTTTAAAGGAATGATGATGATGATGATGATGATGATGATGATGAT
Qy      660  CGATCCAAATTTTACGTTTCCGATGCTTCCGATGCTTCCGATGCTTCCGATGCTTCCGAT
Db      737  CGATCCAAATTTTACGTTTCCGATGCTTCCGATGCTTCCGATGCTTCCGATGCTTCCGAT
Qy      720  AGTCAAGACTCTCTCCGCTGCTCAAGAGCGTGTGTTGACTGTGATAGGTTGAC
Db      797  GGTGAAGAACTCTCTCCGCGGTGCAAGAGCGGTATATGAACTCTGATTAAGATGTTCC
Qy      780  ACCAATCCGATCATTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      857  CCGGTTAATTTATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT

```

QY 840 AGTGGAGGTATACCGAGAGCGGTCTCCCTAAATATTCATTACCTTACGACACTGC 899  
 DB 917 AGGTGCGGGATACCAAGAGCGGCTCCCGAATTTTCACTATCTGACGACCGC 976  
 QY 900 AAGAAACCCACTTGAAGAGTGTGACTTGGGAACCGCTGATGTTCCCTGACTATGC 959  
 DB 977 AAGAAATCTCTCCGACCTAGAT-----AGACCTAATTAAGAGGTAACCTATGCG 1027  
 QY 960 TGGTATGTTATGCTCTGCTTATAGTGCCTTGTATGCTCGCTATTTTGTGAGATTT 1019  
 DB 1028 TGGGATGTTTTCGGGCTCCCGATCAGCCGACTCTGCTGATTTTGGCGGCACTT 1087  
 QY 1020 GCAGATCATATCATGAGAGATACGAGACTGATGCTTACTTACTGCTGCTGCTTGG 1079  
 DB 1088 GCMAATCATCTGCATGAGAGATACGATCCGATGCTTACTTCTGCTGCTGCTGCG 1147  
 QY 1080 AGACTCGACAGAGCCTTGGCATGA 1104  
 DB 1148 AGATTGAGAGAGCCCTGCTTGA 1172

RESULT 11  
 AF038585 1380 bp mRNA linear PLN 15-OCT-1998  
 LOCUS Zea mays pyruvate dehydrogenase kinase isoform 1 mRNA, nuclear gene  
 DEFINITION encoding mitochondrial protein, complete cds.  
 AF038585  
 VERSION AF038585.1 GI:3746430  
 KEYWORDS

ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 1380)  
 Thelen,J.J., Muszynski,M.G., Mierzyk,J.A. and Randall,D.D.  
 Molecular analysis of two pyruvate dehydrogenase kinases from maize  
 J. Biol. Chem. 273 (41), 26618-26623 (1998)  
 MEDLINE 98434573  
 PUBMED 9786901

REFERENCE 2 (bases 1 to 1380)  
 Thelen,J.J., Mierzyk,J.A. and Randall,D.D.  
 Direct Submision  
 Submitted (15-DEC-1997) Biochemistry, University of Missouri, 117  
 Schweitzer Hall, Columbia, MO 65211, USA  
 3 (bases 1 to 1380)

REFERENCE Thelen,J.J., Mierzyk,J.A. and Randall,D.D.  
 Direct Submision  
 Submitted (09-OCT-1998) Biochemistry, University of Missouri, 117  
 Schweitzer Hall, Columbia, MO 65211, USA

REMARK Nucleotide sequence updated by submitter  
 On Oct 15, 1998 this sequence version replaced gi:3695002.  
 Location/Qualifiers  
 1. 1380

ORGANISM="Zea mays"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4577"  
 SS..1146  
 /EC\_number="2.7.1.99"  
 /function="catalyzes the phosphorylation of pyruvate  
 dehydrogenase"  
 /note="PDK1"

/product="pyruvate dehydrogenase kinase isoform 1"  
 /protein\_id="AAC63961.1"  
 /db\_xref="GI:3746431"  
 /translation="MASEPVAVAEVAVRWGMRQTVSLRYMEFGARPTERTLLT  
 AAFPHEKEPIRIARALDLSLPGLSTKPAIKYKDYVVSFEIRSEFPEYRNDK  
 ELAPFMIMIRVHTNVPAIALGVQOLKDLGSKAPPGHEIHQIDLRFYMRI  
 GTRMLGQVVAHDPDPPEGVGLINTKSPPTVARIASEDAKATCMREYSGSPVDI  
 YDPPGFTFYVPHLMLFELVKNSLRAVQERYMSDKLAPVRLIVADGADVTK  
 ISDEGGIPIRSGISRIFTLYISTAEHPDLQHNQVTVAGYGIPIRLVARYFGG

## ORIGIN

DLQIISMEGYTDVAYLHLRLSDSEERLP"

Query Match 46.9%; Score 517.4; DB 8; Length 1380;  
 Best Local Similarity 68.4%; Pred. No. 1,3e-141;  
 Matches 757; Conservative 0; Mismatches 326; Indels 24; Gaps 2;

QY 10 AAGAGGCTAGACGAGATGTTTTGAAAGAGCTTATGAGAGACTTACACAGATGGAGTGC 69  
 DB 52 AGATAGGCTGCGAGAGCGGCTGCGCGGCGGTGGCGAGAGAGTGGCGCGCTGGGCGC 111  
 QY 70 ATGAAACAGACCGGCGGTGAGCCCTCAGTACATGATGAGTGTGGTTTCACTCCACTGAG 129  
 DB 112 ATGCGGCAAGCCGGGTGACGCTGCGCTACATGATGAGTGTGGGCGGCGCCCAAGAG 171  
 QY 130 AAGAACTTCTGATCTCGGCGCAGTTTCTTCAACAGAGCTTCCGATTTGGATCGGAGG 189  
 DB 172 CGGACCTCTCTCTCGCGCGCAGTTCCTGCACAAAGAGCTCCCATTCGCGATCGGCGCA 231  
 QY 190 CGTGCATCGAACTCGAGACGCTGCTTATGCGCTTCTGAGAACTCGCGCTTGAAG 249  
 DB 232 CGGCGCTGACCTCGACCTGCTCCCTTCCGCTTCCACCAAGCCGCTATCTCAAG 291  
 QY 250 GTAAGAGATTGTATGTGAGATCATTCAGGACATGAGACGTTTCTGATCAAGAT 309  
 DB 292 GTGAAAGATTGTATGTGAGATCATTCGCTGAAATCCGGCTTTCCAGAGGTGAGGAGC 351  
 QY 310 ACTGCTGATGAGAAAGAGTTTCAACAGATGATCAAGCTTTAAAGTAAAGGCAACAAC 369  
 DB 352 CAGAAAGATGAGTGCCTTCAACCCAGATGATCAAAATGATCAAGGATGACACACCAAT 411  
 QY 370 GTGATCCCATGATGCTGCGGTGAGACCAAGCTGAGAAAG-----GATG 417  
 DB 412 GTGCTCCCTGATTTGATTTGGGTGACAGCTTAAAGAGACTTGGCGGCTCCAAAG 471  
 QY 418 AAATCTCAAGAAAGCTTATGATGATTCAGTTTCTTGTGCTTCTTCTGCTGCT 477  
 DB 472 GCGTCCCCCGGAATTCATGATGATGATTCAGTTTCTTGTGCTTCTGATGCTCAAG 531  
 QY 478 ATGAGATCCGATGCTTATTCGGGCGAGCATGTTGAGTTGATCAATCAACCAACCACTT 537  
 DB 532 ATGGGATCCGATGCTTATGATGAGGCGAGCTGCTTGTGATGACCTGATCCAGAGCTT 591  
 QY 538 CACACAGTGGTTACATACACACCAAGATGCTCTGATGAGGTGCAAGATGCTAGT 597  
 DB 592 GGTGATAGGGCTTATTAACACAAATGTCAACCAATGATGCTGATGCTGATGCTGAT 651  
 QY 598 GAAGATGCAAGTCAATTTGTTTCAAGAGATGTTGCTTCTCGGAGATTAACATATAT 657  
 DB 652 GAAGATGCAAGTCAATTTGTTTCAAGAGATGTTGCTTCTCGGAGATTAACATATAT 711  
 QY 658 GCGATCCAGTTTCACTTTCCGATGTTTCCGACCACTTGGATTTATGATGATGAG 717  
 DB 712 GGTATCCGGGTTTATCATTTTCAATTTGTTTCAACCAATCTCATGATGATTTTCGAG 771  
 QY 718 TTAGTCAGAACTCTCTCGGTGCTGCAAGAGCGGTTTGTGACTCTATAGGTTGCA 777  
 DB 772 CTGCTGAAATCTCTCTTCCGCGGATGAGAGCAAGAGATGATGATGATGATGATGAT 831  
 QY 778 CCAACATCCGATCATTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 837  
 DB 832 CCTCCAGTTGCAATCATATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 891  
 QY 838 GAAGTGAAGTATCCGAGAGCGGCTCCCTAAATATTTCACTTACCTTACACACACT 897  
 DB 892 GAAGTGAAGTATCCGAGAGCGGCTCCCTAAATATTTTCACTTATATATATATATAT 951  
 QY 898 GCAAGAAACCACTTGAAGAAATGAGACTTGGGAAACCGCTGATGTTCCCTGACTATG 957  
 DB 952 GCAAGAAACCTTCTGATGATGAGAGAG-----TATATAGAGATATAG 999  
 QY 958 GCTGTTATGTTATGCTGCTTATGATGCTTCTGATGCTGCTTATGTTGTTGTTGAGAT 1017

Db	1000	GCTGGAATACGTTAATGGAGATCCCATTAATGCCCTTAACTGTCGCATTAATTCCGTGGGAGC	1057
Oy	1018	TTCGCAATCAATAATCCATGGAGAATACCGGACTGATGCTTAATTGCACTTGTCGTCTT	1077
Db	1060	TTGCAGATCATCTCTTAATGGAGATATGGAACTGATGCTTAACCTTCACTTCAAGCGTG	1115
Oy	1078	GGAGACTGCGCAGAGCTCTTGCCATGA	1104
Db	1120	GGAGATTGGAGAACCTTGCCCTTA	1146
RESULT 12 LOCUS AKI00033			
DEFINITION	Aki00033	1535 bp mRNA linear PLN 24-JUL-2003	
ACCESSION	Aki00033		
KEYWORDS	FLI -CDNA; CAP trapper.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehhartriodeae; Olyzeae; Oryza.		
AUTHORS	1 The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shihiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Otono,Y., Murakami,K., Lida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsumoda,Y., Kuraseki,R., Sugiyama,A., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Marikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., NIKURA,J., Ikeda,T., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusunegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN; Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Kawana,T., Kikuchi,S., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Oato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M., and Hayesbizarzi,Y. Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice		
JOURNAL MEDLINE PUBMED REFERENCE	JOURNAL Science 301 (5611), 376-379 (2003) 27523273 12869764		
AUTHORS	2 (bases 1 to 1535) Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashidume,W., Hayashide,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiroaka,T., Horii,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kuwahara,C., Kuraseki,R., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., MURATA,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., NIKURA,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., OHNO,M., Ohtsuki,K., Oka,M., Ooka,H., Otoe,N., Ota,Y., Otono,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakunuma,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shisshiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tegan-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akashita,S., Tanaka,T., Tomaru,A., Toya,T., Tsumoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasuniishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kamondai, Tsukuba, Ibaraki		

	COMMENT
305-8602, Japan (E-mail:s.kikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007) This clone is one of the 28K full-length cDNA clones from japonica rice.	
URL : http://cdna01.dna.affrc.go.jp/cDNA/ NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yasaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ontenki,K., Shiehiki,T. and Yamamoto,M.	
Fais Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamura,M., Kobayashi,M., Kodama,T., Kuroseki,T., Kusunegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Naitaka,R., Naitaka,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K. Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,U., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Hayashida,K., Hayatsu,N., Hiramaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramaki,T., Hitoko,T., Hori,F., Iida,J., Imamura,K., Imoto,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saichou,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Saseki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Takami-Takeda,Y., Tagawa,A., Takahashi,F., Takahashi-Akhtira,S., Tanaka,T., Tomaru,A., Toyota,T., Waki,K., Yaenishi,A. and Hayashizaki,Y.	
FEATURES	
Source	1..1535 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="mRNA" /cultiyar="Nipponbare" /db_xref="taxon:39947" /clone="J013149J07"
ORIGIN	
Query Match	46.3%; Score 511.4; DB 8; Length 1535;
Best Local Similarity	68.0%; Prid. No. 8e-140;
Matches 754; Conservative	0; Mismatches 331; Indels 24; Gaps 2;
Oy	3 GGCGGTGAAGAAAGCTAGCGAGATGTTTTCGAAGAGCTTGACGAGACTTCACAGATG 62
Db	170 GGGGATGATGCCTGACTCGGAGCCGGCGGCGGAGGGCGGCGTGGGAGAGAGTGGCCGGT 229
Oy	63 GGGATGATTAACCAACGCGGCGTGAAGCTCAGATCATATGATGAGATTGGTTCACTCC 122
Db	230 GGGAGGATGAGGCAACGCGGGGTGACGCTGGCGGTAACAATGAGGAGTTCCGGGCCGCC 289
Oy	123 CACTGAGAGAACCCTTCTGATCTCGGCGCAGTTTCTTCAACAGGAGCTTCGATTCGGAT 182
Db	290 CAAGGAGAGGAACCTCTCGCTCGCGTGGCGGAGTTCTGGCGAGGAGCTCCCATCCGAT 349
Oy	183 CGGAGGCGTTCGATGAACTGAGACGCTGCTTATGGCTTCTTGAAGAACCTGCGGT 242
Db	350 CGCGCGCGCGGCTGACCTGACCTCCCTCCCTTCGCGCTCTCACCAAGCCGCGCAT 409
Oy	243 CTGGAAGTAAAGATTGGTATGTGGAGTCAATTCAGGAGCAATGAGAAGGTTTCTGAGAT 302
Db	410 CTCGAAGTGAAGATTTGGATTGGATCTCAATTCGCGAATTACGTCCTTCGCGAGGT 469
Oy	303 CAAAGTACTGCTGATGAGAAAAGATTCAACAGATGATCAAGGCTGTTAAAGTAGGCA 362
Db	470 GAAGAACCGGAGTATGAGCTTGACATTCACCGAGATGATCAAGATGATTCAGTGGCGCA 529
Oy	363 CAACACGTGTGTTCCATGATGAGCTCTGGGTGTGAACAGCTGAAGAAAG----- 412
Db	530 CAACATGTGTGGCCACCATGCGCTGGGTGTGGCGGAGCTGAAGAAAGCCTGGGTGG 589
Oy	413 --GAATGAACCTCAAGAAAAGCTGATGAGATTGATCATGATTTCTTGATTCGCTTCACTT 470

```

Db      590 CACAAAGCCCTTCCCTCGCATGATGATCATCACTTCCTCGACCGCTTCTACAT
Qy      471 GTCTGTATAGGATCCGTATGCTTACGGGAGACATGTTAGTGCATTAATCCAAACC
Db      650 GTTCAAGATTTGGATCCGATCTGATAGGAGCATGTTTACTTGTGATAGCGCTGATCC
Qy      531 ACCACTTCAACAGTGGGTTACATACACACCAAGATGTCTCTATGAGAGTGGCAAGAA
Db      710 AGAACTGGTGTATTGACTCATTAAGCAAAAGATTATCCCATGTGGTGGCTCAACA
Qy      591 TGTCTATGAAGATGCAAGTCCGTTTGTTCAGAGATATGTTCTGCTCCGAGATAAA
Db      770 TCTTACTGAAGAGACGCTATTTTGACAGAGAAATGATGACAGCCCTGATGTGA
Qy      651 CATATGCGCATCCAAAGTTTCACTTCCGATGTTCCGACCCATTTGACATTTAAGT
Db      830 CATATGAGAGACCCAGATTTTCACTTCCATATGTTAACTACATCTACACTGATGAT
Qy      711 GATAGATTAGTCAAGACTCTCTCGTGTCTCAAGACGGTGTGTTGACTGTGATG
Db      890 GTTTGAATTGTGAAGAAATCCCTCGTGCATGACAGAAAGATATGAACTGTATTA
Qy      771 GGTGACCAACCAATCCGTATCTTGTGCTGATGGAATGGAAGATTTAATAAAGT
Db      950 GATGCGCTCTCTGTTAGATCATAGTCTGATGAGACAGAGACATCAATCAAGAT
Qy      831 CTGATGAAAGTGAAGTATACCGAAGCGGCTCCCTAAATATTAATCACTTACCTTA
Db      1010 TAGTATGAAGTGTGGAATACCAAGAGTGTCTCTGAGAAATTTTCAATATCTTA
Qy      891 CAGCATGCAAGAAACCACTTGAGAAAGATGGAATGGAAGCCGTGATGTTCCCTG
Db      1070 CAGCAAGAGAAACCACTGATCTGATG-----GCCGTAAGAAAGAGAT
Qy      951 GACTATGCTGTGTTATGTTATGCTCTGCTATTAATGCTGATGCTGCTATTTGG
Db      1118 GACCAATGCGCGGATAGGCTATGAGATCCGATTAACCGCTTATGCTGATATTTGG
Qy      1011 TGAAGATTGACATATATCATGAGAGATACGGAGATGATGCTTACTGACCTTGC
Db      1178 CCGGAGCTGCAATATTTCTATGAGAGATATGAGACGATGCTTACTCCTCACTTATC
Qy      1071 TCGCTTGGAGACTCGCAGAGCCTTTGC
Db      1238 ACGGCTAGAGATTGGAGAGGCCCTTGC

```

RESULT 13  
AF323611 1405 bp mRNA linear PLN 08-MAR-2001  
LOCUS Oryza sativa pyruvate dehydrogenase kinase (PDK) mRNA, partial cds.  
DEFINITION AF323611  
ACCESSION AF323611.1 GI:13249141  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Oryza sativa  
Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Eriocaulaceae; Oryzaceae; Oryza.  
REFERENCE  
1 (bases 1 to 1405)  
Peng, Y.-L., Fan, Y.-J. and Zheng, W.-J.  
cDNA cloning and characterization of a pyruvate dehydrogenase  
kinase from rice leaves infected by blast fungus  
Unpublished  
2 (bases 1 to 1405)  
Peng, Y.-L., Fan, Y.-J. and Zheng, W.-J.  
Direct SubMISSION  
Submitted (24-NOV-2000) Plant Pathology Department, China  
Agricultural University, Yuanmingyuan Xilu, Haidian District,  
Beijing 100094, P. R. China  
LOCATION/Qualifiers  
1. 1405  
/organism="Oryza sativa"

```

/mo1_type="mRNA"
/db_xref="taxon:4530"
/note="isolated from leaves infected by blast fungus"
1. 1405
/gene="PDK"
<1. 1032
/gene="PDK"
/codon_start=1
/product="pyruvate dehydrogenase kinase"
/protein_id="AAK16695.1"
/db_xref="GI:13249142"
/translation="QTVSLRYMMEBGRSPTERNLLLSAQFLQKELPIRIARRALE
SLPGLSRKPAIKVNDWLDLDFRIRYPEEVNRNDELAFQIMIRVRRNNVET
MALGVQOLNEDQYTRKIPFADEIHEFLDRYMSRIIGRLIGQVHALDHPBEGV
IGLITBELSPIOGDAASEDARSICLRXGTSVSMRLDYEPPTFFPVVSHLHM
LEIYNLSLAMOYREBYMSDEDVPPRIIVAGGRTIKYSDGGGIPRSGGLPIFYL
YSTAKPPMDPSEBQVIMAGYGYGLPIRLYARIRGDLQIISMEGYGTDAVHLHR
LQDSSEPLP"

```

ORIGIN

Query Match 40.0%; Score 442; DB 8; Length 1405;  
Best Local Similarity 67.7%; Pred. No. 2.9e-119;  
Matches 711; Conservative 0; Mismatches 300; Indels 39; Gaps 5;

```

Qy      76 CAGACGGCGTAGACCTCAGTACATGATGAGATTGCTTCCACTCCACTGAGAAAC
Db      1 CAGAGGGGGTAGACCTGCGTACATGATGAGATTGGGTCGGCCCAAGGCGGAAAC
Qy      136 CTTCGATCTCGCGGAGTTTCTTCAAGAGACTTCCGATTCGATCCGAGCGCTGC
Db      61 CTGCTGCTCTCCGCGAGTTCCTGCAAGAGAGCTCCCAATCCGATCCGCGCCCGCG
Qy      196 ATCGAAGTGAAGCGTGCCTTATGCGCTCTGAGAAACCTGCGCTTGAAGTAA
Db      121 CTGAGCTGATGCTCTCCCTTCCGCTCTCCGAGACCGGCATCTCAAGGTCGG
Qy      256 GATGATGATGAGATCATTCAGGACATGAGACGTTTCTGATCAAGATTAATGCT
Db      181 GATGATGATGAGATCATTCAGGACATGAGATTAATTCAGAGATGAGAAATGAGAC
Qy      316 GATGAGAAAGATTCAACAGATGATCAAGGCTGTTAAAGTAAAGGACAAACGTCGT
Db      241 GACGAGCTGCTTTTAAAGCAGATGATCAAGATGATCAAGTACGCAATATATGTC
Qy      376 CCGATGATGCTGCTGGGTGGAACGAGTGAAGAAAG-----AATGAATC
Db      301 CCAAGATGCACTGGAGATTAACAGCTCAAGAACAGACATATGCAAGAGATC
Qy      424 TAGAAGAAAGCTGATGAGATTCATCATGTTCTTATGCTGCTTATCTGATAGG
Db      361 CCCACTGATTTGAGAGATTCACGAGTTTCTTGAAGATTCATGATCAAGATTTGG
Qy      484 ATCCGTATGCTTATGCGGACATGTTGATGTCATTAATCCAAACCACTTACACA
Db      421 ATCCGATGCTTATGCGGACATGTTGATGTCATTAATCCAAACCACTTACACA
Qy      544 GTGGGTTACATACACCAAGATGCTCTTATGAGAGGTGGCAAGAACTGTGTAAGAT
Db      481 ATAGGCTCATTAATACAGATTTGCCCTTAATACAGTGGTCAAGCTCCAGTGAAGT
Qy      604 GCAAGTCAATTTGTTTCAAGAGATGTTGCT-----CCGAGATTAACATATAT
Db      541 GCCCGTTATTTGTTTGAAGAAATAGGCTCACTAGCTCCGAGATTTGACATCTAT
Qy      658 GCGATCCAAATTTCACTTTCCGATGTTCCGACCCATTTGATTTATGCTATGAG
Db      601 GAGGACCCAACTTTTACATTTCCATATGTTTATCATCATCTCATCTCATGATTTGGA
Qy      718 TTATGCAAGACTCTCT---CCGCTGTCGAAGCGGTTGTTGACTCTGATAGGCT
Db      661 TTGGTGAAGACTCTCTGCGCTGCGCATGACAGGAAAGATATATGATTTCCATGAAGT
Qy      775 GCACCAACCAATCCGATCATTTGTTGCTGATGAGATGAGATGTTAATAAAGTCTCA

```

```

Db      721 GTCTCCAGTTAGAAATATATGCTCTATG-----GGGAGACATATCAAGTTACT 774
Qy      835 GATGAAGGTGAGATATACCGAAGAGCGGTCTCCCTAAATATTCATTACCTCTACAGC 894
Db      775 GATGAAGGTGAGATATACCGAAGAGCGGTCTCCCTAAATATTCATTACCTCTATAGC 834
Qy      895 ACTGCAAGAAACCCACTTGAAGAGATGTGACTTGGGACCGGTGATGTTCCCTGACT 954
Db      835 ACTGCAAGAAATCCACT-----GATATGATTGCTCCCTAGTGAAGAGTAACT 882
Qy      955 ATGCTGTGTTATGTTATGTTGCTGCTATATGCTGTGATGCTGTGCTATTTTGTGA 1014
Db      883 ATGCTGTGTTATGTTATGTTGCTGCTATATGCTGTGATGCTGTGCTATTTTGTGCT 942
Qy      1015 GATTGGAGATCATATCCATGAGAGATACCGGACTGATGCTTACTTTCACCTTGTCTGT 1074
Db      943 GACCTGCAATATCATCTCTATGAGAGATACCGGACCGATGCTTACCTGCACTTGTCAAG 1002
Qy      1075 CTGGAGACTGCGAGAGCCTTTCATGA 1104
Db      1003 TTGGAGATTGAGAGAGCCCTTCCCTTGA 1032

```

```

RESULT 14
LOCUS      ARS66929      899 bp      DNA      linear      PAT 08-OCT-2004
DEFINITION Sequence 965 from patent US 6768041.
ACCESSION  ARS66929
VERSION     ARS66929.1  GI:53983965
KEYWORDS
SOURCE      Unknown.
            Unclassified.
            1 (bases 1 to 899)
REFERENCE   1 (bases 1 to 899)
AUTHORS     Strabala,T. and Nieuwenhuizen,N.
TITLE       Compositions isolated from plant cells and their use in the
            modification of plant cell signaling
            Patent: US 6768041-A 965 27-JUL-2004;
            Location/Qualifiers
            1. 899
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN

```

```

Query Match      29.4%; Score 324.2; DB 6; Length 899;
Best Local Similarity 76.0%; Pred. No. 2.2e-84;
Matches 414; Conservative 0; Mismatches 128; Indels 3; Gaps 1;

Qy      560 CCAAGATGCTCTCTATGAGAGTGGCAAGAAATCTGTGAAGTGCAGAGTGTGTT 619
Db      144 CGAAGATGCTCCAGTTGAGTTGCACTATGCTATTTGAGATGCTCGTCTATTTGTT 203
Qy      620 TCAGAGATATGTTGCTGCTCCGAGATTAACATATATGCGCATCCAAATTTCACTTTC 679
Db      204 TGGGAGATATGCGCACTGCTCCGAGATTAATATATATGCGGATCCCAAGTTTCACTTTC 263
Qy      680 CGATATGTTCCGACCCATTTGCATCTTATGCTGATGAGTATGCAAGAACTCTTCCGCTG 739
Db      264 CGATATGTTCCACACACTTGCATCTTATGATGATTTGATGTTGTTAAGAACTCTTCCGCTG 323
Qy      740 CTGTCAAGAGCGGTTTGTGACTGTGATAGAGTTGCAACCAATCCCTATCATTTGTTG 799
Db      324 CTGTCCAAAGACATTCATGATTCAGATTAAGTATGACCTCTATTTCAATTAATTTGTTG 383
Qy      800 CTGATGAATCGAAGATGTTTACATTAAGATCTCAAGATGAGTGAAGTATATCCAGAA 859
Db      384 CTGATGAAGACGAAGAGTTTACATTAAGATCTCAAGATGAGTGAAGTATATCCAGAA 443
Qy      860 GGGGTCTCCCTAAATTAATTCATCTTACCTCTACAGCACTGCAAGAAACCCACTTGAAGAG 919
Db      444 GTGGTCTTCCCAAAATCTTACGATATCTGTATAGCACTGCGAAGAAACCCCTGTGATGAGA 503

```

```

Qy      920 ATGTGACTTGGAGACCGCTGATGTTCCCTGACTATGCTGTGTTATGTTATGTTCTGC 979
Db      504 ACTCAGATCTTGGAAATAGCTGATAAT---GTGACTATGCGCGGTTATGTTATGGCTTC 560
Qy      980 CTATTAGTGGCTTGTATGCTGCTGCTATTTTGTGAGAAATTTGCAAGATCATATCAATGAAG 1039
Db      561 CAATCAGCGCTGTATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 620
Qy      1040 GATACGGAGCTGATGCTTACTTGCATCTTGTCTGCTTGTGAGACTGCGAGAGCCTTTCG 1099
Db      621 GATATGGAAGTATGATATCTTCAATTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680
Qy      1100 CATGA 1104
Db      681 CGTGA 685

```

```

RESULT 15
LOCUS      ATAC011623
DEFINITION Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence,
            complete sequence.
ACCESSION  AC011623
VERSION     AC011623.5  GI:12408729
KEYWORDS
SOURCE      Arabidopsis thaliana (thale cress)
            Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
            1 (bases 1 to 86022)
            Lin.X., Kaul,S., Town,C.D., Benito,M.-I., Creasy,T.H., Haas,B.,
            Ronning,C.M., Koo,H., Fujii,C.Y., Uteckack,I.R., Batmsstead,M.B.,
            Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.
            Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence
            Unpublished
            2 (bases 1 to 86022)
            Lin.X. and Kaul,S.
            Direct Submission
            Submitted (08-OCT-1999) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA, xlin@tigr.org
            3 (bases 1 to 86022)
            Lin.X.
            Direct Submission
            Submitted (24-JAN-2001) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA
            On Jan 24, 2001 this sequence version replaced gi:12280803.
            Address all correspondence to:
            Xiaoying Lin
            The Institute for Genomic Research
            9712 Medical Center Dr.
            Rockville, MD 20850, USA
            e-mail: xlin@tigr.org
            BAC clone F24P17 is from Arabidopsis chromosome III and is near the
            molecular marker ml403.
            The orientation of the sequence is from SP6 to T7 end of the BAC
            clone.

```

Genes were identified by a combination of three methods: Gene prediction programs including GRL (available by anonymous ftp from arthur.epm.ornl.gov), GeneFinder (Phil Green, University of Washington), GenScan (Chris Burge, <http://www.csb.dtu.dk/netpgene/chrsgene.html>), and NetPlantGene (<http://www.cbs.dtu.dk/netpgene/cbsnetpgene.html>), searches of the complete sequence against a peptide database and the Arabidopsis EST database at TIGR (<http://www.tigr.org/tdb/at/at.html>). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted to be more than two gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy).





Query Match 29.3%; Score 324; DB 8; Length 86022;

Best Local Similarity 64.5%; Pred. No. 5.4e-84;

Matches 725; Conservative 0; Mismatches 95; Indels 304; Gaps 4;

```

QY 224 TCTCGAAGAACTCCCGCTTGAAGTAAAGATTGGTATGTGAGTCATTCAGGACA 283
DB 257 TGTTCATTTTCATGTTTATGTATGTAGTGGCGGATGGTATTTGGAACTTTCAGGGACA 316
QY 284 TGAGAGCGTTTCTGAGTCAAGGATGACTGCGTGAAGAGTGAAGAGTTCACAGATGATCA 343
DB 317 TGAAGGCAATTCCTGAGATTAAAGATTGGGTCAGAGAAAGATTTCACCTAGATGATTA 376
QY 344 AGGCTGTTAAAGTAAGGACAAACACGTGTTCATGATGCTCTGGGTGTGAACGAC 403
DB 377 AGGCTGTCAGAGTAAGCATTAACATGTGTTCATGATGCTCTGGGTGTGAACGAC 436
QY 404 TGAAGAAAGAACTCAACGAAAGCTTGAAGATTCATCAGTTCTTGATGCT 463
DB 437 TCAAGAAAGAACTCAACGAAAGCTTGAAGATTCATCAGTTCTTGATGCT 493
QY 464 TCTACTGTCTGTATAGGATCCGTATGCTTAT----- 497
DB 494 TCTACTGTCTGTATAGGATCCGTATGCTTATGCTTATGCTTATGCTTATGCTTAT 553
QY 498 ----- 497
DB 554 TCAATCTCTTTTCCCTGACTGAAAGTTTGTGCTGATGTGTGTTAATACTTG 613
QY 498 ----- 497
DB 614 TTGCGTTATCATGGAATTTCTCAGGGGACGCTGAGTGTGATTCATTCATTCACCGC 673
QY 536 TTCAACAGTGGGTTACATACACCAAGATGTCTCTATGAGTGGCAAGAAATGCTA 595
DB 674 TTCAATACAGTGGGTTATATACACCAAGATGTCTCTATGAGTGGCAAGAAATGCTA 733
QY 596 GTGAAGATGCAAGTCAATTTGTTTCAGAGATGTGTTCTGCTCCGAGATTAACATAT 655
DB 734 GTGAAGATGCTCGGTCAATTTGTTTCGAGAGTACGGTCTGCAACGGAATTAACATAT 793
QY 656 ATGGCGATCCAGTTTCACTTTTCC----- 680
DB 794 ATGGCGATCCAGTTTCACTTTTCCCGGAGTTCAAGTTGACTGCAATTTTATTTCTTA 853
QY 681 ----- 680
DB 854 GATTCAGTTTGTAGCTTTGTTCTCATGTGTACCAATATCTTTTCTTCTACCTGT 913
QY 681 --GTATGTTCCGACCATTTGCATCTTATGTTATGATGTTCAAGAACTTCTCGCT 738
DB 914 AGGTAATGTTCAACGCAATTTGCACTTATGATGTATGAGTCAAGAACTCTTACCT 973
QY 739 GCTGTCCAGAGCGTGTGTTGCTGCTGATGAGTGGTGCACCAATCCGATCATTTGTT 798
DB 974 GCTGTCCAGAGCGTGTGTTGCTGCTGATGAGTGGTGCACCAATCCGATCATTTGTT 1033
QY 799 GCTGATGGAATCGAAGATGTTACATTA----- 826
DB 1034 GCTGATGGAATCGAAGATGTTACTATTAAGTTGTTACCTGCACTTTTCTCTCTT 1093
QY 827 ----- 826
DB 1094 TTGCTAGCCATTTCTTCTTGTGCTGCTGTTGATCCAAATCATTGTGTAGTTAATA 1153
QY 827 ----AGGTTCAAGATGAGTGAATATACGAGAGCGGTCTCCCTAAATATTTCACT 882
DB 1154 TAACAGGTCTCAGATGAGTGAATATACGAGAGCGGTCTCCCTAAATATTTCACT 1213
QY 883 TACCTCTAAGCACTGCAAGAAACCACTTGAAGAGATGAGTGGAAACGCTGAT 942
DB 1214 TATCTTACAGCACTGCAAGAAACCGCTTGAAGAGATGAGTGGAAATAGCTGAT 1273

```

```

QY 943 GTTCCCTGACTATGCTGTTATGCTATGCTGCTATTTAGTCTGTTATGCTGCC 1002
DB 1274 GTTCCCTGACTATGCTGTTATGCTATGCTGCTATTTAGTCTGTTATGCTGCC 1333
QY 1003 TATTTGGTGAATTTGAGATCATATCCATGAGAGATTAACG 1046
DB 1334 TATTTGGTGAATTTGAGATCATATCCATGAGAGATTAACG 1377

```

Search completed: April 12, 2005, 06:53:52  
Job time : 5153 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2005, 19:30:52 ; Search time 698 Seconds

(without alignments)  
9363.029 Million cell updates/sec

Title: US-10-642-531-1

Perfect score: 1104  
Sequence: 1 atgcgcgcgtgaagaagcctag.....cgcagcagccttgcacatga 1104

Scoring table: IDENTITY\_NIC

Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: N\_Geneseq\_16Dec04:\*  
2: geneseqn1980s:\*  
3: geneseqn1990s:\*  
4: geneseqn2000s:\*  
5: geneseqn2001s:\*  
6: geneseqn2001bs:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1104	100.0	1104	12	ADL71332 Brassica
2	1095.6	99.2	1104	12	ADL71333 Brassica
3	1062.4	96.2	1104	12	ADL71335 Brassica
4	1062	96.2	1104	12	ADL71334 Brassica
5	873.6	79.1	1457	2	AAV49818 A. thaliana
6	524.6	47.5	1602	5	AAV4380 Maize pyr
7	448.2	40.6	1332	5	AAD14379 Maize pyr
8	328.2	29.7	591	13	ACN54458 Maize pyr
9	244.2	22.1	584	13	ADM80480 Maize pyr
10	236.2	21.4	464	3	ACG36807 Arabidops
11	208.2	18.9	371	3	AAV79663 Bucealyptu
12	177.8	16.1	1323	13	ADG49235 Bacterial
13	167.6	15.2	1407	8	ABT21104 Aspergill
14	167.6	15.2	1410	8	ABT19284 Aspergill
15	151.8	13.8	1625	8	ABT20506 Aspergill
16	151.8	13.8	1625	8	ABT18690 Aspergill
17	151.8	13.8	1625	8	ABT19910 Aspergill
18	151.8	13.8	1625	8	ABT18096 Aspergill
19	149.8	13.6	474	13	ACN53319 Cotton an
20	121.4	11.0	236	10	ADB76050 Tomato pl

21	120.2	10.9	1536	6	ABZ32464
22	111.4	10.1	695	13	ADG60834
23	110.2	10.0	285	6	ABT76427
24	108.4	9.8	1476	10	ADB69820
25	80.8	7.3	768	12	ADT40939
26	80.8	7.3	1268	5	AAH80808
27	80.8	7.3	1470	12	ADL61198
28	80.8	7.3	1599	13	ADQ88173
29	80.8	7.3	1599	13	ADR52586
30	79	7.2	1507	10	ADB53519
31	79	7.2	1829	13	ACN39976
32	79	7.2	1866	2	AAH06836
33	71.8	6.5	1798	10	ACA56530
34	71.8	6.5	1798	11	ADP65448
35	71.8	6.5	1798	12	ADT56326
36	71.8	6.5	1798	13	ADG92844
37	71.8	6.5	4889	6	AAH94881
38	71.8	6.5	4889	9	ACH03818
39	70	6.3	466	3	AAH15579
40	66.8	6.1	595	10	ADD33813
41	64.8	5.9	2020	10	ADB69459
42	64.8	5.9	4020	10	ADB69098
43	63.6	5.8	1146	12	ADK71915
44	62.6	5.7	1435	10	ADB58761
45	62.6	5.7	1435	10	ADB53453

#### ALIGNMENTS

##### RESULT 1

ID ADL71332 standard; cDNA; 1104 BP.

XX ADL71332:

XX 20-MAY-2004 (first entry)

XX Brassica napus pyruvate dehydrogenase kinase (PDHK) cDNA.

XX Pyruvate dehydrogenase kinase; PDHK; fatty acid; oil content; flowering;

XX vegetative growth; respiration; transgenic plant; gene therapy;

XX plant protectant; growth regulator; gene; ss.

XX Brassica napus.

XX Key Location/Qualifiers

XX CDS 1..1104

XX /product= "PDHK protein"

XX US2004033606-A1.

XX 19-FEB-2004.

XX 16-AUG-2002; 2002US-00222075.

XX 16-AUG-2002; 2002US-00222075.

XX (ZOU/) ZOU J.

XX (TAYLOR) TAYLOR D C.

XX (MARI) MARI LIA E.

XX ZOU J, Taylor DC, Marillia E;

XX WPI; 2004-246636/23.

XX P-PDB; ADL71336.

XX New purified or recombinant nucleic acid encoding a Brassica pyruvate

XX dehydrogenase kinase protein, for manipulating pyruvate dehydrogenase

XX activity in plants; e.g. seeds with increased or decreased fatty acid and

XX oil content.

ABZ32464 Candida a  
ADG60834 Bacterial  
ABT76427 Corn tass  
ADB69820 C. neofo  
ADT40939 Human kin  
AAH80808 DNA encod  
ADL61198 Human pro  
ADQ88173 Human 177  
ADT52586 Breast ca  
ADB53519 Primary r  
ACN39976 Tumour-as  
AAH06836 Disease a  
ACA56530 Human s1g  
ADP65448 Human pyr  
ADT56326 Human pol  
ADG92844 Pyruvate  
AAH94881 Human DNA  
ACH03818 Human CDN  
AAH15579 Human pro  
ADD33813 Mouse mlt  
ADB69459 C. neofo  
ADB69098 C. neofo  
ADK71915 Human kin  
ADB58761 Toxicity-  
ADB53453 Primary r

PS Claim 2; SEQ ID NO 1; 14bp; English.

CC The present invention relates to novel *Brassica pyruvate dehydrogenase*  
 CC kinase (PDHK) proteins and polynucleotides encoding such proteins.  
 CC Sequences of the invention are useful for manipulating pH activity in  
 CC plants resulting in seeds with increased or decreased fatty acid and oil  
 CC content, plants exhibiting early or delayed flowering time, plants with  
 CC increased or decreased vegetative growth, plants with root systems better  
 CC to withstand low soil temperatures or frost, plants with tissues  
 CC exhibiting higher or lower rates of respiration or plants exhibiting an  
 CC enhanced capacity to accumulate biopolymers which rely on acetyl groups  
 CC as precursors. They are also useful for generating transgenic plants and  
 CC in gene therapy. The present sequence is *Brassica napus pyruvate*  
 CC dehydrogenase kinase (PDHK) cDNA.

CC Sequence 1104 BP; 281 A; 231 C; 294 G; 298 T; 0 U; 0 Other;

Query Match 100.0%; Score 1104; DB 12; Length 1104;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGTGAAGAGCTAGAGAGATGTTTCCGAAGAGCTTGATCGAGAGCTTCAAGA 60  
 DB 1 ATGGCGGTGAAGAGAGCTAGAGAGATGTTTCCGAAGAGCTTGATCGAGAGCTTCAAGA 60  
 QY 61 TGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
 DB 61 TGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
 QY 121 CCCACTGAGAGAACTTTCTGATCTCGGCGGAGTTTCTTCAAGAGAGCTTCCGATTCCG 180  
 DB 121 CCCACTGAGAGAACTTTCTGATCTCGGCGGAGTTTCTTCAAGAGAGCTTCCGATTCCG 180  
 QY 181 ATCGGAGGCGTGGCATGCACTGAGAGCGCTGAGCTTATGAGCTTCTGAGAAAGCTGCC 240  
 DB 181 ATCGGAGGCGTGGCATGCACTGAGAGCGCTGAGCTTATGAGCTTCTGAGAAAGCTGCC 240  
 QY 241 GTCTTGAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 DB 241 GTCTTGAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 QY 301 ATCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
 DB 301 ATCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
 QY 361 CACAAACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
 DB 361 CACAAACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
 QY 421 CTCTAGAGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
 DB 421 CTCTAGAGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
 QY 481 GGGATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
 DB 481 GGGATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
 QY 541 ACAGTGGGTACATACACCAAGATGTCCTATGAGAGGTGCAAGAGATGCTATGAA 600  
 DB 541 ACAGTGGGTACATACACCAAGATGTCCTATGAGAGGTGCAAGAGATGCTATGAA 600  
 QY 601 GATGAGAGTGAATGTTTCAAGAGTATGTTTCTCCGAGATTAACAATATATGAGC 660  
 DB 601 GATGAGAGTGAATGTTTCAAGAGTATGTTTCTCCGAGATTAACAATATATATGAGC 660  
 QY 661 GATCCAGTTTCACTTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
 DB 661 GATCCAGTTTCACTTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
 QY 721 GTCAAGAACTCTCGTCTGCTCAAGAGCGGTTGTTGACTCTGATAGGGTTGCACCA 780  
 DB 721 GTCAAGAACTCTCGTCTGCTCAAGAGCGGTTGTTGACTCTGATAGGGTTGCACCA 780

QY 781 CCAATCCGATCATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
 DB 781 CCAATCCGATCATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
 QY 841 GGTGAGGTATACCGAGAGAGCGTCTCCCTAAATTTCACTTACTTACAGCACTGCA 900  
 DB 841 GGTGAGGTATACCGAGAGAGCGTCTCCCTAAATTTCACTTACTTACAGCACTGCA 900  
 QY 901 AGAACCACCTTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 DB 901 AGAACCACCTTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 QY 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
 DB 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
 QY 1021 CAGATCATATCCATGAGAGATACCGGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
 DB 1021 CAGATCATATCCATGAGAGATACCGGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
 QY 1081 GACTCGCAGAGAGCTTTGGCATGA 1104  
 DB 1081 GACTCGCAGAGAGCTTTGGCATGA 1104

# RESULT 2

ADL71333 standard; cDNA; 1104 BP.

ADL71333;

20-MAY-2004 (first entry)

Brassica rapa pyruvate dehydrogenase kinase (PDHK) cDNA.

Pyruvate dehydrogenase kinase; PDHK; fatty acid; oil content; flowering;

vegetative growth; respiration; transgenic plant; gene therapy;

plant protectant; growth regulator; gene; ss.

Brassica rapa.

Key Location/Qualifiers

FT CDS 1..1104 /tag=a /product="PDHK protein" /trans\_except="(pos:676..678, aa:Phe)

US2004033606-A1.

19-FEB-2004.

16-AUG-2002; 2002US-00222075.

16-AUG-2002; 2002US-00222075.

(ZOUJ/) ZOU J.

(TAYL/) TAYLOR D C.

(MARI/) MARILLIA E.

Zou J, Taylor DC, Marillia E;

WPI; 2004-24636/23.

P-PSDB; ADL71337.

New purified or recombinant nucleic acid encoding a *Brassica pyruvate*

dehydrogenase kinase protein, for manipulating pyruvate dehydrogenase

activity in plants, e.g. seeds with increased or decreased fatty acid and

oil content.

Claim 2; SEQ ID NO 2; 14bp; English.

The present invention relates to novel *Brassica pyruvate dehydrogenase*

kinase (PDHK) proteins and polynucleotides encoding such proteins.

CC Sequences of the invention are useful for manipulating PDH activity in  
 CC plants resulting in seeds with increased or decreased fatty acid and oil  
 CC content, plants exhibiting early or delayed flowering time, plants with  
 CC increased or decreased vegetative growth, plants with root systems better  
 CC to withstand low soil temperatures or frost, plants with tissues  
 CC exhibiting higher or lower rates of respiration or plants exhibiting an  
 CC enhanced capacity to accumulate biopolymers which rely on acetyl groups  
 CC as precursors. They are also useful for generating transgenic plants and  
 CC in gene therapy. The present sequence is *Brassica rapa* pyruvate  
 CC dehydrogenase kinase (PDHK) cDNA.

XX Sequence 1104 BP; 278 A; 231 C; 296 G; 298 T; 0 U; 1 Other;

Query Match 99.2%; Score 1095.6; DB 12; Length 1104;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1098; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGGCGGTGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 60  
 DB 1 ATGGCGGTGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 60  
 QY 61 TGGGATGATGAGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTAC 120  
 DB 61 TGGGATGATGAGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTAC 120  
 QY 121 CCCACTGAGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 180  
 DB 121 CCCACTGAGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 180  
 QY 181 ATGCGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 240  
 DB 181 ATGCGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 240  
 QY 241 GTCCTGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 300  
 DB 241 GTCCTGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 300  
 QY 301 ATCAAGATGATGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTAC 360  
 DB 301 ATCAAGATGATGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTAC 360  
 QY 361 CACAAACAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 420  
 DB 361 CACAAACAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 420  
 QY 421 CTCTAGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 480  
 DB 421 CTCTAGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 480  
 QY 481 GGGATCGATGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTAC 540  
 DB 481 GGGATCGATGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTAC 540  
 QY 541 ACAAGTGGTACATACACACCAAGATGTTTCGAGAGCTTGATCGAGACGTTAC 600  
 DB 541 ACAAGTGGTACATACACACCAAGATGTTTCGAGAGCTTGATCGAGACGTTAC 600  
 QY 601 GATGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 660  
 DB 601 GATGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 660  
 QY 661 GATCCAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 720  
 DB 661 GATCCAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 720  
 QY 721 GTCAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 780  
 DB 721 GTCAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 780  
 QY 781 CCAATCGATGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTAC 840  
 DB 781 CCAATCGATGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTAC 840

QY 841 GGTGAGATGATACGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTAC 900  
 DB 841 GGTGAGATGATACGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTAC 900  
 QY 901 AGAAGCCACTGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTAC 960  
 DB 901 AGAAGCCACTGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTAC 960  
 QY 961 GGTGAGATGATACGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTAC 1020  
 DB 961 GGTGAGATGATACGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTAC 1020  
 QY 1021 CAGATCATGATGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTAC 1080  
 DB 1021 CAGATCATGATGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTAC 1080  
 QY 1081 GACTCGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 1104  
 DB 1081 GACTCGAGAGAGCGGTGAGATGTTTCGAGAGCTTGATCGAGACGTTACAGA 1104

# RESULT 3

ADL71335  
 ID ADL71335 standard; cDNA; 1104 BP.

XX ADL71335;

DT 20-MAY-2004 (first entry)

XX *Brassica carinata* pyruvate dehydrogenase kinase (PDHK) cDNA.

XX Pyruvate dehydrogenase kinase; PDHK; fatty acid; oil content; flowering;

KM vegetative growth; respiration; transgenic plant; gene therapy;

XX plant protectant; growth regulator; gene; ss.

OS *Brassica carinata*.

XX Key Location/Qualifiers

FN CDS 1..1104

FT /tag= a

FT /product= "PDHK protein"

XX US2004033606-A1.

XX 19-FEB-2004.

XX 16-AUG-2002; 2002US-00222075.

XX 16-AUG-2002; 2002US-00222075.

XX (ZOU/J) ZOU J.

XX (TAYLOR D C.) TAYLOR D C.

XX (MARTI/) MARILLIA E.

XX ZOU J, Taylor DC, Marillia E;

XX WPI: 2004-246636/23.

XX P-PSDB; ADL71339.

XX New purified or recombinant nucleic acid encoding a *Brassica pyruvate*

XX dehydrogenase kinase protein, for manipulating pyruvate dehydrogenase

XX activity in plants, e.g. seeds with increased or decreased fatty acid and

XX oil content.

XX Claim 2; SEQ ID NO 4; 14pp; English.

XX The present invention relates to novel *Brassica pyruvate* dehydrogenase

XX kinase (PDHK) proteins and polynucleotides encoding such proteins.

XX Sequences of the invention are useful for manipulating PDH activity in

XX plants resulting in seeds with increased or decreased fatty acid and oil

XX content, plants exhibiting early or delayed flowering time, plants with

XX increased or decreased vegetative growth, plants with root systems better

XX to withstand low soil temperatures or frost, plants with tissues

CC exhibiting higher or lower rates of respiration or plants exhibiting an  
 CC enhanced capacity to accumulate biopolymers which rely on acetyl groups  
 CC as precursors. They are also useful for generating transgenic plants and  
 CC in gene therapy. The present sequence is *Brassica carinata* pyruvate  
 CC dehydrogenase kinase (PDHK) cDNA.

XX Sequence 1104 BP; 287 A; 232 C; 294 G; 291 T; 0 U; 0 Other;

Query Match 96.2%; Score 1062.4; DB 12; Length 1104;  
 Best Local Similarity 97.6%; Pred. No. 0;  
 Matches 1078; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```

QY 1 ATGGCGGTGAAGAGGCTAGCGAGATGTTTTCGAGAGCTTGAAGAGCTTGAAGAGCTTCAACA 60
DB 1 ATGGCGGTGAAGAGGCTAGCGAGATGTTTTCGAGAGCTTGAAGAGCTTGAAGAGCTTCAACA 60
QY 61 TGGGGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB 61 TGGGGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 121 CCCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
DB 121 CCCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 181 ATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 181 ATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 GTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 241 GTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 ATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 301 ATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 361 CACAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 CACAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 CTCTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 CTCTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 GGGATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 GGGATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 ACAATGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 541 ACAATGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 601 GATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 601 GATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 GATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 661 GATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 GTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 721 GTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 CCAATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 CCAATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 GGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB 841 GGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 901 AGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960

```

```

DB 901 AGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 CAGATCATATCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1021 CAGATCATATCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 GACTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1104
DB 1081 GACTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1104

```

```

RESULT 4
ADL71334
ID ADL71334 standard; cDNA; 1104 BP.
XX
AC ADL71334;
XX
DT 20-MAY-2004 (first entry)
XX
DE Brassica oleracea pyruvate dehydrogenase kinase (PDHK) cDNA.
XX
KW Pyruvate dehydrogenase kinase; PDHK; fatty acid; oil content; flowering;
KW vegetative growth; respiration; transgenic plant; gene therapy;
KW plant protectant; growth regulator; gene; sb.
XX
OS Brassica oleracea.
XX
FH Key location/Qualifiers
FT CDS 1..1104
FT
FT /*tag= a
FT /product= "PDHK protein"
FT /transl_except= (pos:253..255, aa:Arg)
FT /transl_except= (pos:280..282, aa:Xaa)
FT /note= "Xaa corresponds to Asn or Asp"
FT /transl_except= (pos:316..318, aa:Asp)
FT /transl_except= (pos:592..594, aa:Ala)
FT /transl_except= (pos:622..624, aa:Arg)
FT /transl_except= (pos:652..654, aa:Ile)
FT /transl_except= (pos:886..888, aa:Leu)
FT /transl_except= (pos:952..954, aa:Thr)
FT /transl_except= (pos:979..981, aa:Pro)
XX
PN US2004033606-A1.
XX
PD 19-FEB-2004.
XX
PF 16-AUG-2002; 2002US-00222075.
XX
PR 16-AUG-2002; 2002US-00222075.
XX
PA (ZOUJ/) ZOU J.
PA (TAYL/) TAYLOR D C.
PA (MARI/) MARILLIA E.
XX
PI Zou J, Taylor DC, Marillia E;
DR WPI; 2004-246636/23.
DR P-PDB; ADL71338.
XX
PT New purified or recombinant nucleic acid encoding a Brassica pyruvate
PT dehydrogenase kinase protein, for manipulating pyruvate dehydrogenase
PT activity in plants, e.g. seeds with increased or decreased fatty acid and
PT oil content.
XX
PS Claim 2; SEQ ID NO 3; 14pp; English.
XX
CC The present invention relates to novel Brassica pyruvate dehydrogenase
CC kinase (PDHK) proteins and polynucleotides encoding such proteins.

```



CC Sequences of the invention are useful for manipulating PDH activity in  
 CC plants resulting in seeds with increased or decreased fatty acid and oil  
 CC content, plants exhibiting early or delayed flowering time, plants with  
 CC increased or decreased vegetative growth, plants with root systems better  
 CC to withstand low soil temperatures or frost, plants with tissues  
 CC exhibiting higher or lower rates of respiration or plants exhibiting an  
 CC enhanced capacity to accumulate biopolymers which rely on acetyl groups  
 CC as precursors. They are also useful for generating transgenic plants and  
 CC in gene therapy. The present sequence is *Brassica oleracea* pyruvate  
 CC dehydrogenase kinase (PDHK) cDNA.

XX Sequence 1104 BP; 286 A; 229 C; 291 G; 289 T; 0 U; 9 Other;

Query Match 96.2%; Score 1062; DB 12; Length 1104;

Best Local Similarity 97.0%; Pred. No. 0;  
 Matches 1071; Conservative 9; Mismatches 24; Indels 0; Gaps 0;

```

QY 1 ATGGCGGTGAAGAGCGTAGCGAGATGTTTTCGAAAGCTTGATCGAGACGTTACAGA 60
DB 1 ATGGCGGTGAAGAGCGTAGCGAGATGTTTTCGAAAGCTTGATCGAGACGTTACAGA 60
QY 61 TGGGAGTGCATGAGAGCAGAGCGGTGAGCGCTCAGGATGATGAGATTCGTTCCACT 120
DB 61 TGGGAGTGCATGAGAGCAGAGCGGTGAGCGCTCAGGATGATGAGATTCGTTCCACT 120
QY 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAAGCTTCGATTCGG 180
DB 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAAGCTTCGATTCGG 180
QY 122 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAAGCTTCGATTCGG 180
DB 122 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAAGCTTCGATTCGG 180
QY 181 ATGCGAGAGCGTCGATGAGAACTCGAGACGCTGCTTATGCGCTTCTGAGAAACCTGCC 240
DB 181 ATGCGAGAGCGTCGATGAGAACTCGAGACGCTGCTTATGCGCTTCTGAGAAACCTGCC 240
QY 181 ATGCGAGAGCGTCGATGAGAACTCGAGACGCTGCTTATGCGCTTCTGAGAAACCTGCC 240
DB 181 ATGCGAGAGCGTCGATGAGAACTCGAGACGCTGCTTATGCGCTTCTGAGAAACCTGCC 240
QY 241 GTCCTAAGGTAGAGATTGCTATGCTGAGTCAATTCAGGAGCAATGAGAGGTTTCCGAG 300
DB 241 GTCCTAAGGTAGAGATTGCTATGCTGAGTCAATTCAGGAGCAATGAGAGGTTTCCGAG 300
QY 241 GTCCTAAGGTAGAGATTGCTATGCTGAGTCAATTCAGGAGCAATGAGAGGTTTCCGAG 300
DB 241 GTCCTAAGGTAGAGATTGCTATGCTGAGTCAATTCAGGAGCAATGAGAGGTTTCCGAG 300
QY 301 ATCAAGATATCTGCTGATGAGAAAGAGTTTCAACAGATGATCAAGGCTGTTAAAGTAAAG 360
DB 301 ATCAAGATATCTGCTGATGAGAAAGAGTTTCAACAGATGATCAAGGCTGTTAAAGTAAAG 360
QY 301 ATCAAGATATCTGCTGATGAGAAAGAGTTTCAACAGATGATCAAGGCTGTTAAAGTAAAG 360
DB 301 ATCAAGATATCTGCTGATGAGAAAGAGTTTCAACAGATGATCAAGGCTGTTAAAGTAAAG 360
QY 361 CACACACACGCTGTTCCATGATGCTCTGGGTGTGAAACCAAGCTGAGAAAGAAATGAAA 420
DB 361 CACACACACGCTGTTCCATGATGCTCTGGGTGTGAAACCAAGCTGAGAAAGAAATGAAA 420
QY 361 CACACACACGCTGTTCCATGATGCTCTGGGTGTGAAACCAAGCTGAGAAAGAAATGAAA 420
DB 361 CACACACACGCTGTTCCATGATGCTCTGGGTGTGAAACCAAGCTGAGAAAGAAATGAAA 420
QY 421 CTCTAAGAAAGGTTGATGATGATCACTTCTTGAATGCTTCTTCTTCTTCTTCTTCTTCT 480
DB 421 CTCTAAGAAAGGTTGATGATGATCACTTCTTGAATGCTTCTTCTTCTTCTTCTTCTTCT 480
QY 421 CTCTAAGAAAGGTTGATGATGATCACTTCTTGAATGCTTCTTCTTCTTCTTCTTCTTCT 480
DB 421 CTCTAAGAAAGGTTGATGATGATCACTTCTTGAATGCTTCTTCTTCTTCTTCTTCTTCT 480
QY 481 GGGATCCGATGCTTATCGGGCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 GGGATCCGATGCTTATCGGGCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 481 GGGATCCGATGCTTATCGGGCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 GGGATCCGATGCTTATCGGGCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 ACAAGTGGTTACATACACACCAAGATGCTCTTATGAGAGGTGCGAAGAAATGCTAGTAA 600
DB 541 ACAAGTGGTTACATACACCAAGATGCTCTTATGAGAGGTGCGAAGAAATGCTAGTAA 600
QY 541 ACAAGTGGTTACATACACCAAGATGCTCTTATGAGAGGTGCGAAGAAATGCTAGTAA 600
DB 541 ACAAGTGGTTACATACACCAAGATGCTCTTATGAGAGGTGCGAAGAAATGCTAGTAA 600
QY 601 GATGCAAGGTCATGTTGTTTTCAGAGATGATGCTCTCGAGAGTAAACCATATATGCGC 660
DB 601 GATGCAAGGTCATGTTGTTTTCAGAGATGATGCTCTCGAGAGTAAACCATATATGCGC 660
QY 601 GATGCAAGGTCATGTTGTTTTCAGAGATGATGCTCTCGAGAGTAAACCATATATGCGC 660
DB 601 GATGCAAGGTCATGTTGTTTTCAGAGATGATGCTCTCGAGAGTAAACCATATATGCGC 660
QY 661 GATGCAAGGTCATGTTGTTTTCAGAGATGATGCTCTCGAGAGTAAACCATATATGCGC 720
DB 661 GATGCAAGGTCATGTTGTTTTCAGAGATGATGCTCTCGAGAGTAAACCATATATGCGC 720
QY 661 GATGCAAGGTCATGTTGTTTTCAGAGATGATGCTCTCGAGAGTAAACCATATATGCGC 720
DB 661 GATGCAAGGTCATGTTGTTTTCAGAGATGATGCTCTCGAGAGTAAACCATATATGCGC 720
QY 721 GTCAAGAACTCTCTCGCTGCTGCTCAAGAGCGGTTGTTGATCTGATGAGGTTGACCA 780
DB 721 GTCAAGAACTCTCTCGCTGCTGCTCAAGAGCGGTTGTTGATCTGATGAGGTTGACCA 780
QY 721 GTCAAGAACTCTCTCGCTGCTGCTCAAGAGCGGTTGTTGATCTGATGAGGTTGACCA 780
DB 721 GTCAAGAACTCTCTCGCTGCTGCTCAAGAGCGGTTGTTGATCTGATGAGGTTGACCA 780
QY 781 CCAATCGATCATGTTGCTGATGAGATGAGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 CCAATCGATCATGTTGCTGATGAGATGAGATGATGATGATGATGATGATGATGATGATGAT 840

```

```

QY 841 GGTGAGATATACCGAGAAAGCGCTCTCCCTAAATAATTCATTACTCTACAGCACTGCA 900
DB 841 GGTGAGATATACCGAGAAAGCGCTCTCCCTAAATAATTCATTACTCTACAGCACTGCA 900
QY 901 AGAAACCCACTTGAAGAGATGTCGACTTGGGAAACCGCTGATGTTCCCTGACTATGCT 960
DB 901 AGAAACCCACTTGAAGAGATGTCGACTTGGGAAACCGCTGATGTTCCCTGACTATGCT 960
QY 901 AGAAACCCACTTGAAGAGATGTCGACTTGGGAAACCGCTGATGTTCCCTGACTATGCT 960
DB 901 AGAAACCCACTTGAAGAGATGTCGACTTGGGAAACCGCTGATGTTCCCTGACTATGCT 960
QY 961 GGTATGATGATGCTGCTGCTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 961 GGTATGATGATGCTGCTGCTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 CAGATCATATCCATGAGAGATACGAGACTGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 CAGATCATATCCATGAGAGATACGAGACTGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 GACTCGCAGAGAGCTTTGCCATGA 1104
DB 1081 GACTCGCAGAGAGCTTTGCCATGA 1104

```

## RESULT 5

AAV49818  
 ID AAV49818 standard; cDNA, 1457 BP.

AAV49818;

02-NOV-1998 (first entry)

A. thaliana PDHK cDNA from clone VAS.

Pyruvate dehydrogenase kinase; PDHK; mitochondrion; transgenic plant;

respiration rate; seed oil content; flowering time; cold resistance;

biomass; biopolymer accumulation; fatty acid synthesis; regulation; ss.

Arabidopsis thaliana.

Key Location/Qualifiers

FT CDS 104..1204 /tag= a /product= "PDHK"

FT /note= "pyruvate dehydrogenase kinase"

MO9835044-A1.

13-AUG-1998.

09-FEB-1998; 98MO-CA000096.

10-FEB-1997; 97US-0038815P.

(CANADA ) NAT RES COUNCIL CANADA.

Taylor DC, Zou J;

WPI; 1998-447241/38.

P-PDB; AAW64724.

claim 1; Fig 2; 86bp; English.

This sequence encodes a mitochondrial pyruvate dehydrogenase kinase,

PDHK, isolated from *Arabidopsis thaliana*. This protein can be used to

produce transgenic plants which have altered respiration rates, seed oil

content and flowering time. Such plants may also have enhanced resistance

to cold, biomass, and capacity to accumulate biopolymers compared to

genetically unmodified plants or seeds. Fatty acid synthesis, vegetative

growth and generation time can also be regulated

Sequence 1457 BP; 395 A; 278 C; 345 G; 439 T; 0 U; 0 Other;

Query Match 79.1%; Score 873.6; DB 2; Length 1457;  
 Best Local Similarity 87.8%; Pred. No. 4.7e-262;  
 Matches 967; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

1 ATGGCGGTGAGAGAGAGCTGAGAGATGTTTTCGAGAGCTTGATGAGAGAGCTTACAGCA 60  
 104 ATGGCAGTGAAGAGAGAGCTTCCGAAATGTTCCGAGAGAGTTGATGAGAGAGTTCACAA 163  
 61 TGGGATGATGAGAGAGAGAGCTGAGAGCTTACAGTATGAGAGAGTTCGACT 120  
 164 TGGGATGATGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 223  
 121 CCAGTGAAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 180  
 224 CCTACTGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 283  
 181 ATGGCAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 240  
 284 GTGGCAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 343  
 241 GTCTGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 300  
 344 GTTGAAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 403  
 301 ATCAAGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 360  
 404 ATTAAGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 463  
 361 CACAAGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 420  
 464 CATACAGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 522  
 421 CTCTAGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 480  
 523 --TTCTGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 580  
 481 GGGATCGGATGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 540  
 581 GGGATCGGATGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 640  
 541 ACAGTGGGATGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 600  
 641 ACAGTGGGATGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 700  
 601 GATGAGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 660  
 701 GATGAGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 760  
 661 GATGAGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 720  
 761 GATGAGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 820  
 721 GTCAGAGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 780  
 821 GTCAGAGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 880  
 781 CCAATCGGATGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 840  
 881 CCAATCGGATGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 940  
 841 GGTGAGAGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 900  
 941 GGTGAGAGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 1000  
 901 AGAAGAGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 960  
 1001 AGAAGAGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 1060  
 961 GGTGAGAGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 1020  
 1061 GGTGAGAGAGAGAGAGAGAGCTTACAGTATGAGAGAGTTCGACT 1120

Query 1021 CAGATCATATCATGAGAGAGATACGAGAGCTTACTGACTTGTCTGCTTGA 1080  
 Db 1121 CAGATCATATCATGAGAGAGATACGAGAGCTTACTGACTTGTCTGCTTGA 1180  
 1081 GACTCCGAGAGAGCTTGTCCATCA 1104  
 1181 GATTCCGAGAGAGCTTGTCCATCA 1204

RESULT 6  
 AAD14380  
 ID AAD14380 standard; cDNA; 1602 BP.  
 AC AAD14380;  
 DT 01-NOV-2001 (first entry)  
 DE Maize pyruvate dehydrogenase kinase (PDK)-2 cDNA.  
 KW Maize; pyruvate dehydrogenase kinase 1; PDK-2; transgenic plant; ss.  
 OS Zea mays.  
 FH Key location/Qualifiers  
 FT CDS /tag= a /product= "Maize PDK-2 protein"  
 FT sig\_peptide /tag= b /product= "Maize PDK-2 protein"  
 FT mat\_peptide /tag= c /product= "Maize PDK-2 protein"  
 FT FT  
 XX US6265636-B1.  
 XX PD 24-JUL-2001.  
 XX PF 15-JUN-1999; 99US-00333423.  
 XX PR 19-JUN-1998; 98US-0089989P.  
 XX PA (PION-) PIONEER HI-BRED INT INC.  
 XX PA (UMOR) UNIV MISSOURI.  
 XX PI Randall D, Thelen J, Miernyk J, Muszynski M, Sewalt V.  
 DR WPI: 2001-463947/50.  
 DR P-PSDB; AAE07838.  
 XX PT Nucleic acids encoding monocot pyruvate dehydrogenase kinase, useful for  
 PT producing plants, e.g. corn cotton and rice with altered characteristics  
 XX e.g. oil content.  
 XX PS Claim 1b; Col 61-64; 36pp; English.  
 XX CC The present sequence is a cDNA coding for maize pyruvate dehydrogenase  
 CC kinase (PDK)-2. The PDK cDNAs are used via standard recombinant DNA  
 CC methodologies for the production of transgenic plants (e.g. corn,  
 CC soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley  
 CC and/or millet) which have altered characteristics and metabolic  
 CC processes, e.g. altered carbon flow into the Krebs cycle and/or  
 CC increased oil content  
 XX SQ Sequence 1602 BP; 388 A; 395 C; 428 G; 391 T; 0 U; 0 Other;

Query Match 47.5%; Score 524.6; DB 5; Length 1602;  
 Best Local Similarity 68.8%; Pred. No. 7.6e-153;  
 Matches 760; Conservative 0; Mismatches 324; Indels 21; Gaps 2;

12 GAAGCTAGAGAGATGTTTTCGAGAGCTTATGAGAGAGCTTACAGAGATGGGATGCAT 71  
 77 GATGCTCGAGAGAGAGAGAGCTTACAGAGAGAGAGAGAGAGCTTACAGAGATGGGATGCAT 136

```

Qy 72 GAAGCAGACGGGCGTGAAGCCTCAGTATCATGATGAGATTGCGTTCACTCCCACTGAG 131
Db 137 GAAGCAGACGGGCGGTTGACCTCGCGGTACATGATGAGATTGCGTTCCGCGCCCAACCG 196
Qy 132 AAACCTTCTGATCTGGCGGCAATTCTTCAAGAGAGCTTCGATTGGATGCGGAGCG 191
Db 197 CAACCTCTCTCTCCCGCGAGTTCTGCAACAGAGAGCTCCCACTCGCTTCGCGCCCG 256
Qy 192 TGGGATGGAATCGAGACGGCTGTATGCGCTCTGAGAAACCTCGCTTGAAGT 251
Db 257 CGGCTGAGCTCACTGCTGCGCTTCGCGCTCTTCAACAGAGCGCGCTCAAGT 316
Qy 252 AAGAGATTGATGATGAGATCATTCAGGACATGAGACGCTTCTGATCAAGATAC 311
Db 317 GCGGACCTGGTACTGATCATTCCTCGGACATCAGATACCTTCTGAGTGAAGACCG 376
Qy 312 TGCTGATGAGAAAGATTACACAGATGATCAAGGCTGTTAAGTAAAGGCAACAAGT 371
Db 377 GAAAGAGAGCTCTCTTCAAGAGATGATCAATGATGATCAAGGCTGCGGCAATGAT 436
Qy 372 GATTCCATGATGCTGCGGTGTAACAGAGTGAAGGATGAACCTCAAGAAA 431
Db 437 GATTCCATGATGCTGCGGTGTAACAGAGTGAAGGATGAAGGATGAAGGATGAAG 496
Qy 432 GCTT-----GATGATTCATCAAGTTCTTGAAGCTTCTTCTTCTCTAT 479
Db 497 GATTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 556
Qy 480 AAGGATCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 539
Db 557 TGGCATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 616
Qy 540 CACAGTGGGTTACATACACAGAGTGTCTCTGATGAGAGTGAAGTGAAGTGAAG 599
Db 617 TGTATAGGCTCTCATACATACAGATGATGATGATGATGATGATGATGATGAT 676
Qy 600 AGATGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
Db 677 AGATGCGCGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 736
Qy 660 CGATCCAGATTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 719
Db 737 CGATCCAGATTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 796
Qy 720 AGTCAAGACCTCTCGGTGTCGCAAGAGCGGTTGATGATGATGATGATGATGAT 779
Db 797 GGTGAAGACCTCTCGGTGTCGCAAGAGCGGTTGATGATGATGATGATGATGAT 856
Qy 780 ACCAATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 839
Db 857 CCGGTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 916
Qy 840 AGTGAAGTATACGAGAGCGGTTCTCTAAATATCACTTACCTTACAGACATG 899
Db 917 AGTGAAGTATACGAGAGCGGTTCTCTAAATATCACTTACCTTACAGACATG 976
Qy 900 AAGAAACCCCACTGAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 959
Db 977 AAGAAACCCCACTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1027
Qy 960 TGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1019
Db 1028 TGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087
Qy 1020 GCAATCATATCCATGAGAGATGAGAGCTGATGATGATGATGATGATGATGATGAT 1079
Db 1088 GCAATCATATCCATGAGAGATGAGAGCTGATGATGATGATGATGATGATGATGAT 1147
Qy 1080 AGACTGAGAGAGCTTGCATGA 1104
Db 1148 AGATTGAGAGAGCTTGCATGA 1172

```

```

RESULT 7
AADI4379
ID AADI4379 standard; cDNA; 1332 BP.
XX
AC AADI4379;
XX
DT 01-NOV-2001 (first entry)
XX
DE Maize pyruvate dehydrogenase kinase (PDK)-1 cDNA.
XX
KW Maize; pyruvate dehydrogenase kinase 1; PDK-1; transgenic plant; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 55..1098
FT FT /*tag= a
FT sig_peptide /product= "Maize PDK-1 protein"
FT FT /*tag= b
FT mat_peptide 109..1095
FT FT /*tag= c
FT FT /product= "Mature maize PDK-1 protein"
XX
XX US6265636-B1.
XX
XX PD 24-JUL-2001.
XX
XX PE 15-JUN-1999; 99US-00333423.
XX
XX PR 19-JUN-1998; 98US-0089998P.
XX
XX PA (PION-) PIONEER HI-BRED INT INC.
XX PA (UMOR ) UNIV MISSOURI.
XX
XX PI Randall DD, Thelen JF, Mierznyk JA, Muszynski MG, Sewalt VH;
XX
XX DR WPI; 2001-463947/50.
XX
XX DR P-PSDB; AAB07837.
XX
XX PT Nucleic acids encoding monocot pyruvate dehydrogenase kinase, useful for
XX PT producing plants, e.g. corn cotton and rice with altered characteristics
XX PT e.g. oil content.
XX
XX PS Claim 1b; Col 57-60; 36pp; English.
XX
XX CC The present sequence is a cDNA coding for maize pyruvate dehydrogenase
XX CC kinase (PDK)-1. The PDK cDNAs are used via standard recombinant DNA
XX CC methodologies for the production of transgenic plants (e.g. corn,
XX CC soybean, sunflower, sorghum, canola, alfalfa, cotton, rice, barley
XX CC and/or millet) which have altered characteristics and metabolic
XX CC processes, e.g. altered carbon flow into the Krebs cycle and/or
XX CC increased oil content
XX
SQ Sequence 1332 BP; 345 A; 311 C; 345 G; 331 T; 0 U; 0 Other;
Query Match 40 6%; Score 448.2; DB 5; Length 1332;
Best Local Similarity 66.1%; Pred. No. 5.5e-129;
Matches 732; Conservative 0; Mismatches 303; Indels 72; Gaps 3;
Qy 10 AAGAGGCTGACGAGATGTTTTCAGAGAGCTTATGAGAGCTTCAAGATGAGGATGC 69
Db 52 AGGATGCGCGTGGAGCGCGGTGCGCGCGGTGCGCGAGAGTGGCGCGCTGGGCGCG 111
Qy 70 ATGAGAGAGCGGCGGTGAGCTTACAGATGATGATGATGATGATGATGATGATGATGAT 129
Db 112 ATGCGGAGAGCGGCGGTGAGCTTACAGATGATGATGATGATGATGATGATGATGATGAT 171
Qy 130 AAGAACTTCTGATCTCGGCGGAGTTCTTCAAGAGAGCTTCCGATGCGGATGCGAG 189
Db 172 CGGACCTCTTCTCGCGGAGGTTCTGACAGAGAGCTCCCAATCCGATGCGGAG 231
Qy 190 CGTGCATGGAATCGAGAGCGTGTGCTTATGCGCTCTGAGAAACCTGCGCTTGAAG 249

```



QY	428	AAAAGCTTATGATGATTCATTCACAGTTCTTGATGCGCTTCCTGCTCGATAGGGATCC	487
Db	62	AGGATCTTGATGATGATTCATCAAGTTCTTCGATCGGTTTACTGTTCACGGATTGGAAATCC	121
QY	488	GTAATGCTTATCGGGCAGCATGTTGAGTTGCATATATCCAAACCAACCATTCACACAGTGG	547
Db	122	GCATGCTTATTTGGGGCAGCAGCTCGACCTGCATTAACCAAAATCCTGCTCCACATGTTGTGG	181
QY	548	GTTACATATACACCAAGATGTCTCCATATGAGAGGTGGCAAGAAATGCTAGTGAAGATGCA	607
Db	182	GCTATATATATATACAAAATATGTTCTCCAGTGGAGGTGGCAAAAATATGCTAGGAGGCTGCAC	241
QY	608	GGTCGATTTGTTTCAGAGAGTATGTTCTGCTCCGAGATPAAATATATATGGCGATCCAA	667
Db	242	GTGCTATTGTTGTGGCGTAAATATGAAAGTGCACCTGATATTATATCTACGGTGATCTCA	301
QY	668	GTTTCACATTTTCCGATATGTTCCGACCCATTTGATCTTATATGCTATATGATATGTCAGA	727
Db	302	GTTTATACATTTCTCTTAAGTTCACAAACACTTGCACATTAATGATATGACACTGCTCAAAA	361
QY	728	ACTCTCTCCGTCGTGTCACAAGCGGGTTTGTGACTCTGATAGAGGTTGCACCAACATCC	787
Db	362	ATTCCTGTGCTGTGCTGCCAAGAGCGCTTATATGATTCAGATTAAGTTTGCACACCTGTTTC	421
QY	788	GTAATCATTTTGTGCTGATGGAATGGAAGATGTTACATTAAGATGGTCTCAGATGAAGGTGGAG	847
Db	422	GAATATATATTTCCCGAGGGAAATCGAAGATGTAACTATTAAGATCTCAGATGAAGGGGGTG	481
QY	848	GTATACCCGAGAGCGGTCTCCCTTAATATATCACTTACCTCTACAGCATGTCAAGAAACC	907
Db	482	GCATTAATAAGAAAGTGGTCTTCCCAAAATTTTCACTTAACCTTTATAGCACTGCCAAAACC	541
QY	908	CACCTGAAAGAAATGTGGACTTGGGAACCGCTGAT	942
Db	542	CCTTGGATAGCATACAGATCTTGGAAACAGATAT	576

RESULT 9
ADM80480
ID ADM80480 standard; DNA; 584 BP.
XX
AC ADM80480;
XX
DT 01-JUL-2004 (first entry)
XX
DE Maize cDNA #1 used in translation initiation codon detection method.
XX
KW Translation initiation codon determination; ATG codon;
KW Quadratic Discriminant Analysis; QDA; initiator codon;
KW pseudoinitiator codon; ATG triplet; bioinformatic; maize; ss.
XX
OS Zea mays.
OS
PN US2004067514-A1.
PN
XX
PD 08-APR-2004.
PD
XX
PF 16-JUL-2003; 2003US-00620796.
PF
XX
PR 17-JUL-2002; 2002US-0396908P.
PR
XX
PA (TABA/) TABASKA J E.
PA
XX
PI Tabaska JE;
PI
XX
DR WP1; 2004-355292/33.
DR
XX
PT Finding translation initiation codons, useful as bioinformatics tool for
PT analyzing files of nucleic acid sequence data, comprises using Quadratic
PT Discriminant Analysis to determine the translation initiation codon.
PT
PS Example 1; SEQ ID NO 1; 19pp; English.
PS
XX

The present invention relates to a method of finding translation initiation codons in a nucleotide sequence using Quadratic Discriminant Analysis (QDA) to determine the translation initiation (ATG) codon in a nucleotide sequence. The method comprises (a) analysing a first data set to measure a combination of features of initiator codons and pseudoinitiator codons and to produce a set of numerical values for the combination of features, (b) evaluating scoring functions by reading a sequence in the vicinity of an ATG triplet and using the scoring functions and the scoring function's parameters to return a numerical score that quantifies how much the ATG triplet resembles an initiator codon, (c) generating a quadratic discriminant function through selection of a combination of feature variables that optimally classifies ATG triplets in a nucleotide sequence as initiator codons or as pseudoinitiator codons based on the output of the scoring functions and through the use of Quadratic Discriminant Analysis, and (d) using the quadratic discriminant function to analyse a second data set of nucleotide sequences by evaluating at least one scoring function for each ATG triplet in the sequences and to calculate the probability of an initiator codon at a position using the output of the analysis. The method is useful for finding translation initiation codons in a nucleotide sequence. The method is useful as a bioinformatics tool for analysing files of nucleic acid sequence data to find translation initiation codons. The present sequence represents a maize cDNA sequence used in the exemplification of the present invention.

Query Match 22.1%; Score 244.2; DB 12; Length 584;

Matches 361; Conservative 0; Mismatches 158; Indels 12; Gaps 1;

QY	12	GAAGGCTAAGCAGATGTTTGTGAAGAGCTTAATCGAGACGTTCAACAATGGGGATATGCAT	71
Db	50	GATGGCGTCGAGACCGCGTGGCCGGGGCGGTGGCGAGAGATGGGCGCTGGGGCAGCAT	109
QY	72	GAAGCAGACGGGGCGTGAACCTCAGGTACATATGATGGATTTGGTTTCACTCCACTGAGAG	131
Db	110	GAAGCAGACGGGGGGTGACCTTGGCGGTATCATATGATGGATTTGGCTCCCGCCACCCAGCG	169
QY	132	AAACCTTCTGATCTTCGGGGCAGTTCTTCAACAAGAGCTTCGATTCGATTCGACAGGGG	191
Db	170	AAACCTGTCTCTTCGCGCAGTTCCTTCAACAAGAGCTCCCATTCGCTTGGCAGCGG	229
QY	192	TGCGATCGAACCTCGAGACGCTGCGCTTATGAGCTCTTGAGAAACCTGCGCTTGAAGGT	251
Db	230	CGCGCTGAGCTCGACCTGCTGCCCTTGGCGCTCTTCAACAAGCCCGCATCTCTCAAGGT	289
QY	252	AAGAGATTGTAATGTGAGATCATTTACAGGACATATGAGACGTTTCTTGATATCAAGATAC	311
Db	290	GCGGGACCTGTAATCTTGGACTATTCCTCGGGAACATCGAATATCTTCCCTGAAGTAGAGACCG	349
QY	312	TGCTGATGAGAAAGATTCACACAGATATCAAGGCTTTAAATGTAAGGACAAACAAGCT	371
Db	350	GAAGCAGACGCTCGCTTTTCAAGCAGATATCAATATGGTCAAGATGGCGCATTAACAATGT	409
QY	372	GCTTCCCATGATGCGCTCTGGGTGTGAACCAAGCTGAAGAAAGGAATGAATCTTACGAAAA	431
Db	410	GCTTCCCAACCATGGCTCTTGGGAGTGCACAGCACTGAAGAGAGCGTGGCGGCTTCACAGAA	469
QY	432	GCTT-----GATGAGATTCATCAGTTTCTTATGCGCTCTAATTGTCTCGTAT	479
Db	470	GCTTCCATTCGAATGCATATGATATGACGAGATTCCTTGAACCGGTTCTTAATATTCAAAGAA	529
QY	480	AGGATCCGTAATGCTTATCGGGCAGCATGTTGATGCTCAATATCCAAACC	530
Db	530	TGGCATTCGCAATGCTGATAGGGCAGCATATGCGCTTGTGATATGACCTCTTAACC	580

RESULT 10	
AAC36807	
ID	AAC36807 standard; DNA; 464 BP
XX	
AC	AAC36807;

XX 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 15144.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
XX protein identification; signal transduction pathway; metabolic pathway;  
XX promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
XX 05-MAR-1999; 99US-0123180P.  
XX 09-MAR-1999; 99US-0123548P.  
XX 23-MAR-1999; 99US-0125788P.  
XX 25-MAR-1999; 99US-0126264P.  
XX 29-MAR-1999; 99US-0126785P.  
XX 01-APR-1999; 99US-0127462P.  
XX 06-APR-1999; 99US-0128234P.  
XX 08-APR-1999; 99US-0128714P.  
XX 16-APR-1999; 99US-0129845P.  
XX 19-APR-1999; 99US-0130077P.  
XX 21-APR-1999; 99US-0130449P.  
XX 23-APR-1999; 99US-0130510P.  
XX 23-APR-1999; 99US-0130891P.  
XX 28-APR-1999; 99US-0131449P.  
XX 30-APR-1999; 99US-0132048P.  
XX 04-MAY-1999; 99US-0132407P.  
XX 05-MAY-1999; 99US-0132484P.  
XX 06-MAY-1999; 99US-0132485P.  
XX 06-MAY-1999; 99US-0132487P.  
XX 07-MAY-1999; 99US-0132487P.  
XX 11-MAY-1999; 99US-0132863P.  
XX 14-MAY-1999; 99US-0134218P.  
XX 14-MAY-1999; 99US-0134219P.  
XX 14-MAY-1999; 99US-0134221P.  
XX 14-MAY-1999; 99US-0134370P.  
XX 18-MAY-1999; 99US-0134376P.  
XX 19-MAY-1999; 99US-0134941P.  
XX 20-MAY-1999; 99US-0135124P.  
XX 21-MAY-1999; 99US-0135353P.  
XX 24-MAY-1999; 99US-0135629P.  
XX 25-MAY-1999; 99US-0136021P.  
XX 27-MAY-1999; 99US-0136382P.  
XX 28-MAY-1999; 99US-0136782P.  
XX 01-JUN-1999; 99US-0137222P.  
XX 03-JUN-1999; 99US-0137528P.  
XX 04-JUN-1999; 99US-0137502P.  
XX 07-JUN-1999; 99US-0137724P.  
XX 08-JUN-1999; 99US-0138094P.  
XX 10-JUN-1999; 99US-0138540P.  
XX 10-JUN-1999; 99US-0138847P.  
XX 14-JUN-1999; 99US-0139119P.  
XX 16-JUN-1999; 99US-0139452P.  
XX 16-JUN-1999; 99US-0139453P.  
XX 17-JUN-1999; 99US-0139492P.  
XX 18-JUN-1999; 99US-0139454P.  
XX 18-JUN-1999; 99US-0139455P.  
XX 18-JUN-1999; 99US-0139456P.  
XX 18-JUN-1999; 99US-0139457P.  
XX 18-JUN-1999; 99US-0139458P.  
XX 18-JUN-1999; 99US-0139459P.  
XX 18-JUN-1999; 99US-0139460P.  
XX 18-JUN-1999; 99US-0139461P.  
XX 18-JUN-1999; 99US-0139462P.  
XX 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140033P.  
PR 23-JUN-1999; 99US-0140035P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 28-JUN-1999; 99US-0140919P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 06-JUL-1999; 99US-0142055P.  
PR 08-JUL-1999; 99US-0142390P.  
PR 09-JUL-1999; 99US-0142803P.  
PR 12-JUL-1999; 99US-0142920P.  
PR 13-JUL-1999; 99US-0142977P.  
PR 14-JUL-1999; 99US-0143542P.  
PR 15-JUL-1999; 99US-0143624P.  
PR 16-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 19-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145224P.  
PR 27-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146385P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 04-AUG-1999; 99US-0147038P.  
PR 05-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148555P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149175P.  
PR 17-AUG-1999; 99US-0149368P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 18-AUG-1999; 99US-0149428P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 25-AUG-1999; 99US-0149930P.  
PR 26-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.



```

PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      21.4%; Score 236.2; DB 3; Length 464;
Best Local Similarity 83.5%; Pred. No. 6,7e-63;
Matches 268; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1 ATGCGCGTGAAGAAAGCTTGCAGATCTTTCCGAAGAGCTTGTTCAGAGAGCTTCCACAGA 60
DB 144 ATGCGCGTGAAGAAAGCTTGCAGATCTTTCCGAAGAGCTTGTTCAGAGAGCTTCCACAAA 203
QY 61 TGGGGATGCATGAAGAGAGCGGCGTGCAGCTCAGAGTACATGATGAGATTCCGTTCCACT 120
DB 204 TGGGGTTCATGAAGAGAAACCGGTGTTAGCTTACATGATGAGATTCCGTTCCAAA 263
QY 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAACAAGAGCTTCCGATTCCG 180
DB 264 CCACTGAGAGAGAAATCTTTGATTTCTGCTCAGTTTTCATTAAGAGCTTCCGATTCCG 323
QY 181 ATGCGAGAGCGTGCAGTGCAGACTCGAGAGCGCTTATGAGCGCTTCTGTGAAGAACTGCC 240
DB 324 GTGCGCAGAGAGAGCGAGTGCAGATTCAGAGCGCTTCTTATGAGTCTTCTGTAAACCTGCC 383
QY 241 GTCTTAAGAGTAAGATTGATGTGAGTCAATTCAGGAGACATGAGAGCGTTTCTCGAG 300

```

```

DB 384 GTTTAGAGCTGCTGATTTGATTTGGAATCTTTCAGGAGACATGAGAGATTTCCGAG 443
QY 301 ATCAGAGATCTCTGATGAG 321
DB 444 ATTAAGATTCGGGTGACGAG 464

RESULT 11
AAA79663 standard; cDNA; 371 BP.
ID AAA79663;
AC AAA79663;
XX
XX 27-NOV-2000 (first entry)
DT
XX Eucalyptus grandis cell signalling involved polynucleotide SEQ ID NO:464.
DE
XX Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
KM plant cell signalling; modulation; transgenic plant; pathogen; growth;
KM environmental change; development; cell proliferation; differentiation;
XX elongation; survival; disease resistance; nutrient metabolism; ss.
XX Eucalyptus grandis.
OS
XX
XX WO200042171-A1.
PN
XX 20-JUL-2000.
PD
XX 11-JAN-2000; 2000WO-US000724.
PF
XX
XX 12-JAN-1999; 99US-00228986.
PR 01-NOV-1999; 99US-0162866P.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX Strabala TJ, Nieuwenhuizen NJ;
XX WPI; 2000-476052/41.
XX
XX Isolated polynucleotide encoding a polypeptide involved in cell signaling
PT used for generating transgenic plants with modified responses to external
PT signals.
XX
XX Claim 1; Page 208; 527pp; English.
PS
XX AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
CC pine (Pinus radiata also known as Monterey pine). The protein sequences
CC are involved in cell signalling. The polynucleotide and protein sequences
CC can be used to modify the response of plant cells to external signals
CC e.g. environmental changes or pathogens during the growth and development
CC of a plant. They can be used to modify cell proliferation,
CC differentiation, elongation and survival, resistance to disease and
CC nutrient metabolism. Examples of modifications which can be produced are
CC altered fruit ripening and senescence of leaves and flowers e.g. to delay
CC senescence and prolong the life of cut flowers or enhance senescence of
CC reproductive organs to engineer sterile plants. Other modifications can
CC be used to delay senescence in selected cell types or organs providing
CC fruit and vegetables which have a longer shelf life between harvest and
CC consumption, or to decrease branching frequency in forest tree species
CC giving long stretches of valuable knot-free clear wood which can be used
CC in solid timber furniture and veneers
XX
XX Sequence 371 BP; 97 A; 74 C; 93 G; 107 T; 0 U; 0 Other;
SQ

Query Match      18.9%; Score 208.2; DB 3; Length 371;
Best Local Similarity 76.8%; Pred. No. 3,4e-54;
Matches 268; Conservative 0; Mismatches 78; Indels 3; Gaps 1;

QY 697 TTGCATCTTATGAGTATGATGATTAAGTCAAGAACTCTTCGCGTGTGTCCAGAGCGGTTT 756
DB 2 TTGCATCTAATGATATTTAGATTGTTAAGAACTCTTTCGCGTGTGTCCAGAGAGGATTC 61

```

The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source. The invention also relates to a transformed plant comprising the recombinant DNA construct and a method of producing a transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant

-----



OS Aspergillus fumigatus.  
 XX WO200286090-A2.  
 XX 31-OCT-2002.  
 XX  
 XX 23-APR-2002; 2002WO-US013142.  
 XX  
 XX 23-APR-2001; 2001US-0285697P.  
 XX 27-APR-2001; 2001US-0287066P.  
 PR 05-JUN-2001; 2001US-0295890P.  
 PR 09-JUL-2001; 2001US-0303899P.  
 PR 31-AUG-2001; 2001US-0316362P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM,  
 PI WPI; 2003-093124/08.  
 DR  
 XX  
 PT New purified or isolated nucleic acids of essential genes of Aspergillus  
 PT fumigatus, useful for treating or preventing infections by A. fumigatus,  
 PT or for treating a non-infectious disease in a subject e.g. cancer.  
 XX  
 PS Disclosure; Page; 175pp; English.  
 XX  
 CC The invention relates to novel purified or isolated nucleic acids of  
 CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of  
 CC the invention are used to treat or prevent infections by a pathogenic  
 CC organism such as A. fumigatus, to treat a non-infectious disease in a  
 CC subject (e.g. cancer), to prevent or inhibit formation of an object  
 CC by A. fumigatus, or to prevent or inhibit formation on a surface of a  
 CC biofilm comprising A. fumigatus. The polynucleotides are useful for  
 CC expressing recombinant protein for characterization, screening or  
 CC therapeutic use, as markers for host tissues in which the pathogenic  
 CC organisms invade or reside, for comparing with the DNA sequence of A.  
 CC fumigatus to identify duplicated genes or paralogues having the same or  
 CC similar biochemical activity and/or function, for comparing with DNA  
 CC sequences of other related or distant pathogenic organisms to identify  
 CC potential orthologous essential or virulence genes, for selecting and  
 CC making oligomers for attachment to a nucleic acid array for examination  
 CC of expression patterns, for raising anti-protein antibodies, as an  
 CC antigen to raise anti-DNA antibodies or to elicit another immune  
 CC response, and for identifying polynucleotides encoding the other protein  
 CC with which binding occurs or to identify inhibitors of the binding  
 CC interaction. The polypeptides may be used to raise antibodies or to  
 CC elicit immune response, as a reagent in assays designed to quantitatively  
 CC determine levels of the protein in biological fluids, as a marker for  
 CC host tissues in which pathogenic organism invade or reside, and to  
 CC isolate correlative receptors or ligands in the case of virulence  
 CC factors. This polynucleotide sequence represents one of the essential  
 CC genes of Aspergillus fumigatus of the invention  
 XX  
 SQ Sequence 1410 BP; 354 A; 377 C; 350 G; 329 T; 0 U; 0 Other;  
 Query Match 15.2%; Score 167.6; DB 8; Length 1410;  
 Best Local Similarity 52.6%; Pred. No. 3.5e-41;  
 Matches 418; Conservative 0; Mismatches 364; Indels 12; Gaps 2;  
 QY 323 AAGAGTTACACAGATGATCAAGCTGTTAAAGTAAGCAACAACGTGTTCCATGA 382  
 DB 617 AACGGTTTGCAGAGAGCTGCACATATCAACGCGACACGACGTCGTGACCAACG 676  
 QY 383 TGGCTCTGGGTGGAACCAAGTGAAGAAAGATGAATCTACGAAAAGCTTGATGAGA 442  
 DB 677 TAGCCCAAGGCAATTTGGAGATGGAAGCGAAGCCAAAGTTTACAAATGATCCACCG 736  
 QY 443 TTCATCAGTTTCTTGAATGCTTACTGTTCTGTAAGGATCCGATATGCTTATCGGCG 502  
 DB 737 TCCAAATCGTTCCAGATCGCTTTATATGTCGCCGATTTGATAGCAATGTTAAATCGGTC 796  
 QY 503 AGCATTTGAGTTG-----CATATCCAAACCCACCACTTCACACAGTGGTTATCATAC 556

DB 797 AACATATGCCCTGAGGAAACAGACGATGTTGCCACCCGAACTACGTGGTATCATCT 856  
 QY 557 AACCAAGATGATCTCTATGATGAGTGAAGAAATGCTGATGAAGTGCATTT 616  
 DB 857 GCAACAAAGAAATGTCAGAGAGTGCCTCGAGGCCATTAACAAATGCTGTTGCTCT 916  
 QY 617 GTTTCAGAGATATGTTCT-----GTCGCGAGATTAACATATATGCGATCCAAAGTT 670  
 DB 917 GTAGGACTACTATGATGATCTTTTTCGAAGCACCGAATGATCAGCTCATCTCAGACGACC 976  
 QY 671 TCACTTTTCGTAATGTTCCGACCATTTTGCATTTATGATGATGATGATGATGATGATGAT 730  
 DB 977 TCACTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1036  
 QY 731 CTCTCGTGTCTGCAAGACGGTGTGTTGATCTGATGATGATGATGATGATGATGATGATGAT 790  
 DB 1037 CGCTGCGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1096  
 QY 791 TCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 850  
 DB 1097 TCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1156  
 QY 851 TACCGAAGACGGTCTCTCTAAATATTCATCTTACCTTACAGACGCTGCAAGAAACGAC 910  
 DB 1157 TACCCGTTCTTCATCCCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1216  
 QY 911 TTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 970  
 DB 1217 CCATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1276  
 QY 971 ATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1030  
 DB 1277 AGGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1336  
 QY 1031 CCATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1090  
 DB 1337 GTATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1396  
 QY 1091 AGCCTTGCCATGA 1104  
 DB 1397 AGCCCTTCATGA 1410  
 RESULT 15  
 ID ABR20506  
 ID ABR20506 standard; DNA; 1625 BP.  
 AC ABR20506;  
 XX  
 DT 16-APR-2003 (first entry)  
 XX  
 DE Aspergillus fumigatus essential gene #2864.  
 XX  
 KM Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;  
 KM cancer; contamination; biofilm; antibody; immune response; ds.  
 OS Aspergillus fumigatus.  
 XX  
 PN WO200286090-A2.  
 PD 31-OCT-2002.  
 XX  
 PF 23-APR-2002; 2002WO-US013142.  
 XX  
 PR 23-APR-2001; 2001US-0285697P.  
 PR 27-APR-2001; 2001US-0287066P.  
 PR 05-JUN-2001; 2001US-0295890P.  
 PR 09-JUL-2001; 2001US-0303899P.  
 PR 31-AUG-2001; 2001US-0316362P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM,

XX  
DR WPI, 2003-093124/08.  
XX  
PT New purified or isolated nucleic acids of essential genes of *Aspergillus*  
PT fumigatus, useful for treating or preventing infections by *A. fumigatus*,  
PT or for treating a non-infectious disease in a subject e.g. cancer.  
XX  
PS Disclosure; Page; 175pp; English.  
XX  
CC The invention relates to novel purified or isolated nucleic acids of  
CC essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of  
CC the invention are used to treat or prevent infections by a pathogenic  
CC organism such as *A. fumigatus*, to treat a non-infectious disease in a  
CC subject (e.g. cancer), to prevent or contain contamination of an object  
CC by *A. fumigatus*, or to prevent or inhibit formation on a surface of a  
CC biofilm comprising *A. fumigatus*. The polynucleotides are useful for  
CC expressing recombinant protein for characterization, screening or  
CC therapeutic use, as markers for host tissues in which the pathogenic  
CC organisms invade or reside, for comparing with the DNA sequence of *A.*  
CC *fumigatus* to identify duplicated genes or paralogues having the same or  
CC similar biochemical activity and/or function, for comparing with DNA  
CC sequences of other related or distant pathogenic organisms to identify  
CC potential orthologous essential or virulence genes, for selecting and  
CC making oligomers for attachment to a nucleic acid array for examination  
CC of expression patterns, for raising anti-protein antibodies, as an  
CC antigen to raise anti-DNA antibodies or to elicit another immune  
CC response, and for identifying polynucleotides encoding the other protein  
CC with which binding occurs or to identify inhibitors of the binding  
CC interaction. The polypeptides may be used to raise antibodies or to  
CC elicit immune response, as a reagent in assays designed to quantitatively  
CC determine levels of the protein in biological fluids, as a marker for  
CC host tissues in which pathogenic organism invade or reside, and to  
CC isolate correlative receptors or ligands in the case of virulence  
CC factors. This polynucleotide sequence represents one of the essential  
CC genes of *Aspergillus fumigatus* of the invention  
XX  
SQ Sequence 1625 BP; 398 A; 432 C; 401 G; 394 T; 0 U; 0 Other;

Query Match 13.81; Score 151.8; DB 8; Length 1625;  
Best Local Similarity 52.54; Pred. No. 3.3e-36;  
Matches 386; Conservative 0; Mismatches 337; Indels 12; Gaps 2;

QY 323 AAGAGTTCAACAGATGATCAAGCTGTTAAAGTAAGGACAAACAAGTGTCCATGA 382  
DB 765 AAGGTTTGGAAGAGCTGCACATATCAAACGGGACACGACGCTGTGACACCG 824  
QY 383 TGGCTCTGGGTGAAACGAGCTGAAGAAAGAAATGAATCTACGAAAAGCTTGATGAGA 442  
DB 825 TAGCCAGGGCATTTGGAATGGAAGCGAAACGCCAAGCTTAAACAATGATTCACCG 884  
QY 443 TTCATAGTTCTTGATCGCTTCTACTGTCTGTATAGGATCCGTAATGCTTATGGGGC 502  
DB 885 TCCATCGTTCCTAGATCGCTTTTATATGTCGGTATGTGATACGAATGTATATCGTC 944  
QY 503 AGCATGTTGAGTTG-----CATATCCAAACCCACCATTCACACAGTGGGTTACATAC 556  
DB 945 AACACATCGCCCTGAGGAAACAGACGATTTGGCCACCGAATCTACGTGATATCATCT 1004  
QY 557 ACACCAAGATGTTCTCTATGAGGTGGCAAGAAATGCTAGTGAAGATGCAAGTCAATTT 616  
DB 1005 GCACCAAGACAAATGTCGGAAGAGTCCCTCGAGGCAATTGACAAATGCTGTTGCTCT 1064  
QY 617 GTTTCAGAGATGAGTCT-----GTCGGAGATAAACAATATAGGCAATCCAAAGTT 670  
DB 1065 GTAGAGACTAGTATGGCTTTTGAAAGACACCAAGATTACCTCATCTGCAAGAGACACC 1124  
QY 671 TCACCTTTCCGATGTTCCGACCATTTGATCTTATAGTGTATGAGTTAGTCAAGAACT 730  
DB 1125 TCAACTTCATGATGATGTTCCGCTGACCTGTACACATGCTCTTGAAGACTTAAAGATT 1184  
QY 731 CTCTCCGTGCTGTCCAAAGACGGTGTGATCTGTATAGGTTGACACCAATCCGTA 790  
DB 1185 CGCTGGTGTCTGTGAGAGACCAACGGCGCGGACAAAGAAAGCAATCCCTGTGACAAAGG 1244

QY 791 TCATGTGCTGATGGAATGGAATGTTATCAATTAAGGTCGAGATGAAGTGAAGTA 850  
DB 1245 TCATTTGTGCAAGAGAAAGAGACATCACCTTCAAAATCTCCGATGAAGGTGTGCA 1304  
QY 851 TACCGAAGACGGCTCTCCCTAAATATTAATCACTTACCTTACAGCACTGCAAGAAACCCAC 910  
DB 1305 TACCCGTTCTTCATCCGTTGCGTGTGACGTACATGTACACTACGCTGATCAACCGC 1364  
QY 911 TTGAAGAATGTGACTTGGAAACCGCTGATGTTCCCTGATAGCTGTATGCTT 970  
DB 1365 CCAATGTGATCCGATTCGATTAAGACGACTTCAAGCTCCATAGGCTGGGCTTGGGT 1424  
QY 971 ATGCTGCTTATTAGTCTGTTATGCTGCTGCTAATTTGGTGGAGATTGCAATCATAT 1030  
DB 1425 ACGGATTCGAATCACTGCTGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1484  
QY 1031 CCATGGAAGATACG 1045  
DB 1485 GTATGGAAGGTTAAG 1499

Search completed: April 12, 2005, 05:23:03  
Job time : 704 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2005, 23:08:37 ; Search time 227 Seconds  
(without alignments)  
7957.927 Million cell updates/sec

Title: US-10-642-531-1

Perfect score: 1104  
Sequence: 1 atgcgcgtcgaagaagcctag.....cgcaggagccttgcacatga 1104

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/prodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/prodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	873.6	79.1	1457	US-09-355-912A-1	Sequence 1, Appl
2	873.6	79.1	1457	US-10-202-428-1	Sequence 1, Appl
3	524.6	47.5	1602	US-09-333-423-3	Sequence 3, Appl
4	448.2	40.6	1332	US-09-333-423-1	Sequence 1, Appl
5	324.2	29.4	899	US-10-101-464A-965	Sequence 965, App
6	208.2	18.9	371	US-10-101-464A-464	Sequence 464, App
7	120.2	10.9	624	US-09-248-796A-4156	Sequence 4156, App
8	80.8	7.2	1601	US-09-949-016-5850	Sequence 5850, App
9	79	7.2	1866	US-08-878-989-13	Sequence 13, Appl
10	79	7.2	1866	US-09-272-796-13	Sequence 13, Appl
11	71.8	6.5	1797	US-09-949-016-4569	Sequence 4569, App
12	71.8	6.5	1797	US-09-949-016-4569	Sequence 4569, App
13	62.4	5.7	1422	US-09-949-016-3858	Sequence 3858, App
14	60.8	5.5	1422	US-09-949-016-377	Sequence 377, App
15	60.8	5.5	72992	US-09-949-016-17592	Sequence 17592, A
16	54.6	4.9	18864	US-09-949-016-12119	Sequence 12119, A
17	54.6	4.9	18864	US-09-949-016-15598	Sequence 15598, A
18	53.4	4.8	601	US-09-949-016-163093	Sequence 163093, A
19	53.4	4.8	601	US-09-949-016-163094	Sequence 163094, A
20	53.4	4.8	15220	US-09-949-016-16311	Sequence 16311, A
21	45.4	4.1	738	US-09-848-708B-22	Sequence 218B, App
22	39.6	3.6	1141	US-08-425-299A-3	Sequence 22, Appl
23	38	3.4	3176	US-08-486-663A-14	Sequence 14, Appl
24	38	3.4	3176	US-08-247-804B-13	Sequence 13, Appl
25	38	3.4	3176	US-08-767-942A-14	Sequence 14, Appl
26	38	3.4	3176	US-09-070-060-10	Sequence 10, Appl
27	38	3.4	3177	US-09-070-060-10	Sequence 10, Appl

28	38	3.4	3177	US-09-357-746-10	Sequence 10, Appl
29	38	3.4	7218	US-08-232-463-14	Sequence 14, Appl
30	37	3.4	2817	US-09-949-016-1916	Sequence 1916, App
31	37	3.4	3063	US-08-294-312B-3	Sequence 3, Appl
32	37	3.4	3063	US-08-468-024B-3	Sequence 3, Appl
33	37	3.4	3063	US-09-708-200-12	Sequence 12, Appl
34	37	3.4	3063	US-09-023-655-1435	Sequence 1435, App
35	37	3.4	3063	US-08-465-679-3	Sequence 3, Appl
36	37	3.4	3063	US-09-788-657-8	Sequence 8, Appl
37	37	3.4	3063	US-09-712-691-10	Sequence 10, Appl
38	37	3.4	3063	US-09-707-468C-10	Sequence 10, Appl
39	36.6	3.3	154746	US-09-827-688-8	Sequence 8, Appl
40	34.8	3.2	231	US-09-248-796A-7339	Sequence 7339, App
41	34.8	3.2	3177	US-09-826-312A-1	Sequence 1, Appl
42	34.8	3.2	3177	US-09-542-497A-1	Sequence 1, Appl
43	34.4	3.1	431	US-09-513-999C-23703	Sequence 23703, A
44	34.2	3.1	2235	US-09-153-804-2	Sequence 2, Appl
45	34.2	3.1	3079	US-09-643-597-116	Sequence 116, App

## ALIGNMENTS

RESULT 1									
US-09-355-912A-1									
; Sequence 1, Application US/09355912A									
; Patent No. 6500670									
; GENERAL INFORMATION:									
; APPLICANT: Zou, Jitao									
; FILE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE									
; TITLE REFERENCE: 40942-1									
; CURRENT APPLICATION NUMBER: US/09/355,912A									
; PRIOR FILING DATE: 1999-10-15									
; PRIOR APPLICATION NUMBER: PCT/CA98/00096									
; PRIOR FILING DATE: 1998-02-09									
; NUMBER OF SEQ ID NOS: 9									
; SOFTWARE: Patent version 3.0									
; SEQ ID NO 1									
; LENGTH: 1457									
; TYPE: DNA									
; ORGANISM: Arabidopsis thaliana									
US-09-355-912A-1									
Query Match									
Best Local Similarity 87.6%; Pred. No. 9.6e-293;									
Matches 967; Conservative 0; Mismatches 134; Indels 3; Gaps 1;									
QY	1	ATGCGCGTGAAGAGCTTGCAGATGTTTGGAAAGCTGATCGAGGACGTTCCAGA	60						
DB	104	ATGCGCGTGAAGAGCTTGCAGATGTTTGGAAAGCTGATCGAGGACGTTCCAGA	163						
QY	61	TGGGATGATGAGAGAGCGGCGCTGAGCTCAGGTATGATGAGGAGCTTCCACT	120						
DB	164	TGGGATGATGAGAGAGCGGCGCTGAGCTCAGGTATGATGAGGAGCTTCCACT	223						
QY	121	CCGACTGAGAGAACTTCTGATCTGGCGCAGTTCTTCAAGAGCTTCCGATTGG	180						
DB	224	CCGACTGAGAGAACTTCTGATCTGGCGCAGTTCTTCAAGAGCTTCCGATTGG	283						
QY	181	ATGCGCGTGAAGAGCTTGCAGATGTTTGGAAAGCTGATCGAGGACGTTCCAGA	240						
DB	284	ATGCGCGTGAAGAGCTTGCAGATGTTTGGAAAGCTGATCGAGGACGTTCCAGA	343						
QY	241	GTTTGAAGATGAGATGTTGATGAGTATTCAGGAGACATGAGAGCTTCTCTGAG	300						
DB	344	GTTTGAAGATGAGATGTTGATGAGTATTCAGGAGACATGAGAGCTTCTCTGAG	403						
QY	301	ATCAGAGATGCTGCTGATGAGAAAGCTTCAACAGATGATCAAGGCTGTTAAGTAA	360						
DB	404	ATCAGAGATGCTGCTGATGAGAAAGCTTCAACAGATGATCAAGGCTGTTAAGTAA	463						

```

OY 361 CACAACAGTGGTTCCTCCATGATGCTCTGGGTGACACACTGAGAGAAATGAA 420
Db 464 CATACAGATGGTTCCTCCATGATGCTCTGGGTGACACACTGAGAGAAATGAA 522
OY 421 CTCTACGAAAAGCTTGATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 480
Db 523 --TTCTGGAATCTTGATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 580
OY 481 GGGATCCGATGATTCATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 540
Db 581 GGGATCCGATGATTCATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 640
OY 541 ACAGTGGTTCATACACACCAAGATGCTCTGATGATGATTCATGATTCCTGATTCAT 600
Db 641 ACAGTGGTTCATACACACCAAGATGCTCTGATGATGATTCATGATTCCTGATTCAT 700
OY 601 GATGCAAGTTCATGATTCATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 660
Db 701 GATGCTCGATCAATTTGTTCTCGAAGATGATGATTCATGATTCCTGATTCATGATTCAT 760
OY 661 GATCCAGTTCATGATTCATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 720
Db 761 GATCCAGTTCATGATTCATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 820
OY 721 GTCAAGAACTCTCTCCGTGCTGCTCCAGAGCGGTTTGTGATCTGATGATGATTCATGATTCAT 780
Db 821 GTCAAGAACTCTCTCCGTGCTGCTCCAGAGCGGTTTGTGATCTGATGATGATTCATGATTCAT 880
OY 781 CCAATCCGATTCATGATTCATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 840
Db 881 CCAATCCGATTCATGATTCATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 940
OY 841 GGTGAGGTATACCGAAGAGCGGTCCTCTAAATATTCATCTGATTCATGATTCCTGATTCAT 900
Db 941 GGTGAGGTATACCGAAGAGCGGTCCTCTAAATATTCATCTGATTCATGATTCCTGATTCAT 1000
OY 901 AGAACCACCTTGAAAGAGATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 960
Db 1001 AGAACCACCTTGAAAGAGATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 1060
OY 961 GGTATGATTCATGATTCATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 1020
Db 1061 GGTATGATTCATGATTCATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 1120
OY 1021 CAGATCATTCATGATGAGATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 1080
Db 1121 CAGATCATTCATGATGAGATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 1180
OY 1081 GACTCGCAGAGCGCTTGCCATGA 1104
Db 1181 GATTCCGAAAGACCTTACCCTGA 1204

```

```

RESULT 2
US-10-202-428-1
; Sequence 1, Application US/10202428
; Patent No. 6825039
; GENERAL INFORMATION:
; APPLICANT: Zou, Jitao
; APPLICANT: Taylor, David
; TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE
; FILE REFERENCE: 40942-1
; CURRENT APPLICATION NUMBER: US/10/202,428
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/05/355,912
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/CA98/00096
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/038,815
; PRIOR FILING DATE: 1997-02-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0

```

```

; SEQ ID NO 1
; LENGTH: 1457
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-202-428-1

Query Match 79.1%; Score 873.6; DB 4; Length 1457;
Best Local Similarity 87.6%; Pred. No. 9,6e-293;
Matches 967; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

OY 1 ATGCGCGTGAAGAAAGCTGACGATGTTTGAAGAGCTGATGAGACCTTACACA 60
Db 104 ATGCGAGTGAAGAAAGCTGACGAAATGTTCCGAAGATTTAATCAAGATTTACAAA 163
OY 61 TGGGGATGATGAAGACGACGCGGTGACCTCAGGTACATGATGATGATTCGATTCAC 120
Db 164 TGGGGATGATGAAGACGACGCGGTGATCCTTAATATATATGATGATGATTCGATTCAAA 223
OY 121 CCCACTGAGAAACCTTCTGATTCGCGGCGATTTCTTCAAGAGACCTTCGATTCG 180
Db 224 CTTACTGAGAGAAATCTTGTGATTCGCTCAGTTTGTGATGATGATGATTCGATTCGATTCG 283
OY 181 ATGCGAGGCGTGCATGACGATTCGACGCTGCTTATGCGCTCTGAGAAACCTGCG 240
Db 284 GTGCGCAGAGAGCATGACGATTCGACGACGCTTCTTATGCTCTTCTGATTAACCTGCG 343
OY 241 GTCTTGAAGTAAAGATTTGATGATGATTCATGATTCAGGACATGAGACGTTCTGAG 300
Db 344 GTTTGAAAGTCCGGATTTGATTTGAAATCTTCAAGGACATGAGACATTCCTGAG 403
OY 301 ATCAAGATGATGATGAGAAAGTTCAACAGATGATCAAGGCTGTTAAAGTAAAG 360
Db 404 ATTAAGATTCGGGATGACGAGAAAGATTTCACTGATGATTAAGGCTGTTAAAGTAAAG 463
OY 361 CACAACAGTGGTTCCTCCATGATGATTCGATGATTCATGATTCCTGATTCATGATTCAT 420
Db 464 CATACAGATGGTTCCTCCATGATGATTCGATGATTCCTGATTCATGATTCCTGATTCAT 522
OY 421 CTCTACGAAAAGCTTGATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 480
Db 523 --TTCTGGAATCTTGATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 580
OY 481 GGGATCCGATGATTCATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 540
Db 581 GGGATCCGATGATTCATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 640
OY 541 ACAGTGGTTCATACACACCAAGATGCTCTGATGATGATTCATGATTCCTGATTCAT 600
Db 641 ACAGTGGTTCATACACACCAAGATGCTCTGATGATGATTCATGATTCCTGATTCAT 700
OY 601 GATGCAAGTTCATGATTCATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 660
Db 701 GATGCTCGATCAATTTGTTCTCGAAGATGATGATTCATGATTCCTGATTCATGATTCAT 760
OY 661 GATCCAGTTCATGATTCATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 720
Db 761 GATCCAGTTCATGATTCATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 820
OY 721 GTCAAGAACTCTCTCCGTGCTGCTCCAGAGCGGTTTGTGATCTGATGATGATTCATGATTCAT 780
Db 821 GTCAAGAACTCTCTCCGTGCTGCTCCAGAGCGGTTTGTGATCTGATGATGATTCATGATTCAT 880
OY 781 CCAATCCGATTCATGATTCATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 840
Db 881 CCAATCCGATTCATGATTCATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 940
OY 841 GGTGAGGTATACCGAAGAGCGGTCCTCTAAATATTCATCTGATTCATGATTCCTGATTCAT 900
Db 941 GGTGAGGTATACCGAAGAGCGGTCCTCTAAATATTCATCTGATTCATGATTCCTGATTCAT 1000
OY 901 AGAACCACCTTGAAAGAGATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 960
Db 1001 AGAACCACCTTGAAAGAGATGATGATTCATGATTCCTGATTCATGATTCCTGATTCAT 1060

```



FEATURE:  
NAME/KEY: CDS  
LOCATION: (55)...(1095)  
US-09-333-423-1

Query Match 40.6%; Score 448.2; DB 3; Length 1332;  
Best Local Similarity 66.1%; Pred. No. 1e-144;  
Matches 732; Conservative 0; Mismatches 303; Indels 72; Gaps 3;

```

Oy 10 AAGAGGCTAGCGAGATGTTTTTCGAAGCTTGATCGAGACGTTTCAAGATGGGAGATC 69
Db 52 AGAGTGGCTCGAGAGCCGAGCGCGGCGGCGGAGAGAGTGGCGCGCTGGGCGCG 111
Oy 70 ATGAGCAGACGCGGCTGAGGCTCAGATACATGATGAGTTCCGTTCCATCCCATGAG 129
Db 112 ATGCGCCAGACCGGGGTCAGCTCGCTACATGATGAGTTCCGGGGCGCGCCACAGAG 171
Oy 130 AGAAACCTTCTGATTCGGCGCAGTTTTCACAAAGAGCTTCCGATTGAGTGGAGG 189
Db 172 CGGACCTCTCTTCGCGCGCAGTCTCTGCAAGAGAGCTCCCATCCGATCGCGCA 231
Oy 190 CGTGCATCGAATCGAGACGCTGCTTATGAGCTCTCTGAGAAAGCTCCGCTTGAAG 249
Db 232 CGGGCGCTCGACCTCGACTCGCTCCCTCGGCTCTTCCACCAAGCCCGCATCTCAG 291
Oy 250 GTAAAGATTGTTATGTTGAGTCAATTCAGGAGCATGAGAGCGTTTCTGAGATCAAGAT 309
Db 292 GTGAAGATTGTTATGTTGAGTCAATTCGTAATCCGCTCTTCCAGAGGTGAGGAC 351
Oy 310 ACTGCTGAGAGAAAGTTTACACAGATGATGATCAAGCTGTAAAGTAAAGCACAAC 369
Db 352 CAGAAAGTAAAGCTCGCTTACCAAGATGATGATCAAGAGTAAAGTAAAGTAAAGTAAAG 411
Oy 370 GTGTTCCCATGATGAGTCTGAGTGAACCAAGCTGAGAAAG-----GAGT 417
Db 412 GTGTTCCCATGATGAGTCTGAGTGAACCAAGCTGAGAAAG-----GAGT 417
Oy 418 AAATCTTACGAAAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 477
Db 472 GCGTTCCTCCCGGAAATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 531
Oy 478 ATAGGATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 537
Db 532 ATAGGATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 538
Oy 538 CACACAGTGGTTACATACACCAAGATGATGATGATGATGATGATGATGATGATGATGAT 597
Db 559 -----ATAACACAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 603
Oy 598 GAGATGCAAGTGCATGTTGTTTCAAGAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 657
Db 604 GAGATGCAAGTGCATGTTGTTTCAAGAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 663
Oy 658 GCGATTCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 717
Db 664 GCGATTCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 723
Oy 718 TTAGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 777
Db 724 CTGTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 783
Oy 778 CCAACATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837
Db 784 CCAACATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 843
Oy 838 GAGATGCAAGTGCATGTTGTTTCAAGAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 897
Db 844 GAGATGCAAGTGCATGTTGTTTCAAGAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 903
Oy 898 GCAAGAAACCACTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 957
Db 904 GCAAGAAACCACTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 951

```

```

Oy 958 GCTGTTATGTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1017
Db 952 GCTGTTATGTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1011
Oy 1018 TTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077
Db 1012 TTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1071
Oy 1078 GAGACTCGAGAGGCTTTCATGA 1104
Db 1072 GAGACTCGAGAGGCTTTCATGA 1098

```

## RESULT 5

```

US-10-101-464A-965
; Sequence 965, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and their use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10101,464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 965
; LENGTH: 899
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-965

```

Query Match 29.4%; Score 324.2; DB 4; Length 899;  
Best Local Similarity 76.0%; Pred. No. 1.1e-101;  
Matches 414; Conservative 0; Mismatches 128; Indels 3; Gaps 1;

```

Oy 560 CCAAGATGTTCTTATGAGTGGGCAAGATGCTTAAAGATGCAAGTGGATTTGTT 619
Db 144 CCAAGATGTTCTTATGAGTGGGCAAGATGCTTAAAGATGCAAGTGGATTTGTT 203
Oy 620 TCAGAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 679
Db 204 TCAGAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 263
Oy 680 GGTATGTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 739
Db 264 GGTATGTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 323
Oy 740 CTGTCGAAGAGCGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 799
Db 324 CTGTCGAAGAGCGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 383
Oy 800 CTGATGAATGAAGATGTTTCAATTAAGGTTCTCAGATGAGAGTGAAGTATACCGAGAA 859
Db 384 CTGATGAGAGAGAGATGTTTCAATTAAGGTTCTCAGATGAGAGGAGTGAAGTATACCGAGAA 443
Oy 860 GCGGTTCCCTTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 919
Db 444 GCGGTTCCCTTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 503
Oy 920 ATGAGACTTGGGAAACGCTGATGTTCCCTGATGATGATGATGATGATGATGATGATGATGAT 979
Db 504 ATGAGACTTGGGAAACGCTGATGTTCCCTGATGATGATGATGATGATGATGATGATGATGAT 560

```



US-09-949-016-5850

Query Match 7.3%; Score 80.8; DB 4; Length 1601;  
 Best Local Similarity 54.5%; Pred. No. 8.5e-17;  
 Matches 216; Conservative 0; Mismatches 162; Indels 18; Gaps 2;

QY 682 TATGTTCCGACCCATTTTCATCTTATGTTATGAGTACAGAACTCTCCGTCCT 741  
 DB 729 TATGTCCTCACAATCTTTTCAATATGCTATTTGATGTTCCAGAACTCAATAGAGCG 788  
 QY 742 GTCCAAAGCGCGTTGTTGACTCTGATAGGGTTGCACCAATCCGATCATTTGCT 801  
 DB 789 ACAGTTGACTCTGTGAAGACAGAAAGAGGCTACCTGCTGTAAACCTCGTACT 848  
 QY 802 GATGGAATCGAAGATGTACATAAAGCTCAGATGAAGTGAAGATCCAGAAC 861  
 DB 849 TTGGGTAAAGAACTTATCCATTAAGTACAGTACCTGAGTGGTGGTCCACTTGA 908  
 QY 862 GGTCTCCCTAAATATTCATCTTACCTTACAGACACTGCAGAAACCACTGAAGAAGAT 921  
 DB 909 AAAATAGATCGTCTTTTAACTACATGATTTCTATCTGCTCTACACCA----- 957  
 QY 922 GTGACTTGGGACCCGCTGATGTCCTCTGACTATGCTGTTATGTTATGCTGCT 981  
 DB 958 ---GCTGGAGCCCTACACAGAGCTCCCTT---TGGCTGATTTGGTTATGTTGCCA 1010  
 QY 982 ATTAGTCTGTATGCTGCTATTTTGTGAGAGATTGAGATGATATGATGATGAGGA 1041  
 DB 1011 ATTTCGCGTCTGTATGCTAGATATTTTCAAGAGAACTGTAAACTGTATTCATGGAAGGA 1070  
 QY 1042 TACGGACTGATGCTTACTTGCACCTGTCTGCTCT 1077  
 DB 1071 GTGGTACTGATGCTGCTATTTATTTGAAGGCTCT 1106

## RESULT 9

US-08-878-989-13  
 Sequence 13, Application US/08878989  
 Patent No. 5885803

## GENERAL INFORMATION:

APPLICANT: Bandman, Olga  
 APPLICANT: Hillman, Jennifer L.  
 APPLICANT: Corley, Neil C.  
 APPLICANT: Guejler, Karl G.  
 APPLICANT: Lai, Preeti  
 APPLICANT: Goli, Surya K.  
 APPLICANT: Shah, Purvi  
 TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Drive  
 CITY: Palo Alto  
 STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: PASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/878,989

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J J

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0321 US

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555  
 TELEFAX: 415-845-4166

## INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 1866 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: COLANTU03  
 CLONE: 1340712  
 US-08-878-989-13

Query Match 7.2%; Score 79; DB 2; Length 1866;  
 Best Local Similarity 45.5%; Pred. No. 4e-16;  
 Matches 372; Conservative 0; Mismatches 430; Indels 15; Gaps 2;

QY 83 GCGTAGCCTCAGATCATGATGAGTTCGTTCCACTCCACTGAGAGAACTTTGTA 142  
 DB 470 GCCTAACGCCACCATGATGCTTACGCTGGCCGCTCTGAGAGCGGACCTTCTGA 529  
 QY 143 TCTCGCGAGATTCTTTCACAAGAGCTTCCGATTCGATCGCGAGGGGTGGATCGAAC 202  
 DB 530 AAAGTGTCTGTAACCTGCAAGAACTTCCAGTGAATTTCTACCCGATCAAGGCT 589  
 QY 203 TCGAGAGCTGCTTATGAGCTCTCTGAGAAACCTGCGCTTGAAGTTAGATTGCT 262  
 DB 590 TCGCTGCTCTCTTTCATCATTTGCTGCAACCCACCACTACTGACGTCATGAGTAT 649  
 QY 263 ATGTGAGTCAATTCAGGGGACATGAGAGGCTTCTCGATATCAAGATATCTGTATGGA 322  
 DB 650 ATATCGTCTCTTCCAGAACTGACAGACTTCTCCATCAAGAACAGGCGGACGAG 709  
 QY 323 AAGAGTTCAACAGATGATCAAGGCTGTTAAAGTAAGGACAAACGTTGCCATGA 382  
 DB 710 CCAGATACGACAGGCTGGTGGAGAGCTGCTGATGACCAAGATGTGTGACCTCT 769  
 QY 383 TGGCTCTGGGTGTAACCACTGAAAGAAAGATAAACTTCAGAAAGCTTGATGAGA 442  
 DB 770 TGGCAGAGGG-----CTTACGTGAGAGCGGAGACATGAGATGAAAGCTCG 820  
 QY 443 TTCATCAGTTCTTATGCTTCTTACTTGTCTGATATGAGGATCCGATATGCTTATGCGGC 502  
 DB 821 TCCGCTACTTCTGACAAAGAGCTGACTTGAAGCTTGAATTCGCAATGTGGCAGCG 880  
 QY 503 AGCATGTTGATGATTAATCCAAACCCACCACTTACACAGTGGTTATACACCA 562  
 DB 881 ATCACCTGGCGCTGATAGAGCAAGCC-----TGACTTGTGGCATCATCTGATCTC 934  
 QY 563 AGATGTTCTTATGAGAGTGGGAGAGATGCTAGTGAAGATGCAAGTGCATTTGTTCA 622  
 DB 935 GTCTCTCAACCAAGAAATTAATGGAAGTGGTGAATTTGCCAGAGCGCTGTGAGC 994  
 QY 623 GAGAGTATGTTCTCTCCGAGATAAATACATATATGCGATCAAGTTTCACTTTTCGCT 682  
 DB 995 ACAATGATGCAATGCGCCGCTGTCGCCATCAATGCGCATGTGCTCCGCTTCCCT 1054  
 QY 683 ATGTTCCAGCCCATTTGATCTTATGTTGATGATGATGATGATGATGATGATGATGAT 742  
 DB 1055 TCATCCCTATGCACTGATACATCTCGCGAGAGCTCTCAAGAAATCCATGAGAGCA 1114  
 QY 743 TCCAAAGAGGTTTGTATCTGTATAGGTTGACCAACCAATCCGATCATTTGTGTG 802  
 DB 1115 CAATGAGAGCACTTACACTCTCCATCAATATGCCAGATGTGTATCCATTCGCA 1174  
 QY 803 ATGAAATCAGAGATGTTCAATTAAGTCTCAGATGAGTGAAGTATACCGAAGCG 862  
 DB 1175 ACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1234  
 QY 863 GTCTCCCTAAATATTTACTTACTTACAGACTGC 899





```

OY 154 TTTCCTCACAAGAGGCTCCGATTCGGATCGGAGCGGCGATCGAACTCGAGCGCTG 213
DB 388 TTTTTCGCAAGAAATTCCTGAGACTGCGCAACATTCGAAAGAAATGATATCTC 447
OY 214 CCTTATGCGCTCTCGAGAAACCTGCGCTCTGAAAGTAAGATTTGGATTTGAGTCA 273
DB 448 CCGACCCATTTAGTAATACCTCTTCAAGTCAATTTGTTAAAGTGATATACAGC 507
OY 274 TTCAGGAGCATAGAGCGCTTCCTGATCA-----AGATATCTGCTGATGAGAAAG 327
DB 508 CTGATGAGATTTGGTGAATTCATAGAAAGCCAGATGACCAAGAAAGCATTTATCAAC 567
OY 328 TTCACACAGATGATCAAGGCTGTTAAAGTAAAGGCAACAAGTGTCCATGAGTGGCT 387
DB 568 TTTGTAGATACACTCATCAAGTTGGAATAGACCAATATGTATGCTCCATACAGCA 627
OY 388 CTGGGTGTGAACCACTGAAAGAAAGAAATGAAATCTTAC---GAAAGCTTGATGAGAT 444
DB 628 CAAAGAAATCATAGAGTATTAAGATGCTGTACAGTTGACCAAGTCAACATCAAAATCTT 687
OY 445 CATCAGTTCTTGATCGCTCTACTGTCTGTATAGGATCCGATGCTTATCGGCGAG 504
DB 688 CAATATTTCTTGATCGATTTTATCACTGACCGTATTTCTACTCGAGTGTGATGAACGAG 747
OY 505 CATGTTGAGTTGCATATCCAAACCCACACTTCAACA-----GTGGGTTACATACAC 558
DB 748 CACATTTCTTATATTAGTACCTCAGACAGAGAAACCAACCAATTGGAAGATGAT 807
OY 559 ACCAAGATGTCTCCATAGAGTGTGCAAGAAATGCTAGTAAAGTGAAGTGCAGTTGT 618
DB 808 CTTAAGTGTATGTGTGATGAGTGTGCAAGATGCTTGTAGTGTTCAGAGATGCTGT 867
OY 619 TTCAGAGATGTGTTCTGCTCCGAGATGAA-----CATATATGCG 660
DB 868 GATCAGATTTATTTATCATCTCCAGATTTAAAGTTACACAGAAATGGAATTTTCCA 927
OY 661 GATCCAGTTTCACTTTTCCGTATGTTCCGACCATTTGCAATTTATGTTATGAGTTA 720
DB 928 GACCAACCAATTCACATCGTGTATGTTCTTCACTCATCATATGATCTTTGAACTA 987
OY 721 GTCAAGAACTCTCTCCGTGCTGCTCAAGAGCGGTTTGTATCTGTATAGGTTGACCA 780
DB 988 TTTTAAAGATGAATGCGGGGCAAGTTGAAACACAGAAATCAGCTTCTT---ACA 1044
OY 781 CCAATCCGTATCATTTGCTGATGATGAAAGATGTTAAATTAAGTCTCAGATGAA 840
DB 1045 CCAATAGAGGTATTTGTTGCTTGGGAAAGAAAGACCTTACATTAAGATTTCAAGACA 1104
OY 841 GGTGAGGTATTAACGAGAAAGCGCTCCCTAAATAATTCCTTACCTTACAGACTGCA 900
DB 1105 GAGAGTGTGTTCCCTGAGAAATTTATGACCGCTCTTTATGATATACATCTCAGTGA 1164
OY 901 AGAAACCACTTGAAGAAAGTGTGACTTGGGAAACCGCTGATGTTCCCTGATGAGCT 960
DB 1165 CCAACGCTGTGATGATTAATCCCGAATG-----CTCTTTGGCT 1206
OY 961 GGTATGATTAATGCTGCTTATGATGCTTATGCTGCTTATTTGGTGAAGATTTG 1020
DB 1207 GGTATGATTAATGCTGCTTATGATGCTTATGCTGCTTATTTCAAGAGATTTG 1266
OY 1021 CAGATCATATCCATGAGAGTACGGAGCATGATCTTACCTTGCACTT 1067
DB 1267 AATCTCTACTCTTATCAGATATGGAACGATGCTATCATCTT 1313

```

RESULT 12  
 US-09-016-434-1128  
 ; Sequence 1128, Application US/09016434  
 ; Patent No. 6500938  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Janice Au-Young  
 ; APPLICANT: Jeffrey J. Seilhamer  
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

```

/ / TITLE OF INVENTION: PATHWAY GENE EXPRESSION
/ / NUMBER OF SEQUENCES: 1490
/ / CORRESPONDENCE ADDRESS:
/ / ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
/ / STREET: 3174 PORTER DRIVE
/ / CITY: PAUL AUTO
/ / STATE: CALIFORNIA
/ / COUNTRY: USA
/ / ZIP: 94304
/ / COMPUTER READABLE FORM:
/ / MEDIUM TYPE: Floppy disk
/ / OPERATING SYSTEM: IBM PC compatible
/ / SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ / CURRENT APPLICATION DATA:
/ / APPLICATION NUMBER: US/09/016,434
/ / FILING DATE: HEREMITH
/ / CLASSIFICATION:
/ / PRIOR APPLICATION DATA:
/ / APPLICATION NUMBER:
/ / FILING DATE:
/ / CLASSIFICATION:
/ / ATTORNEY/AGENT INFORMATION:
/ / NAME: Zeller, Karen J.
/ / REGISTRATION NUMBER: PA-0002 US
/ / TELEPHONE: (650) 855-0555
/ / TELEFAX: (650) 845-4166
/ / INFORMATION FOR SEO ID NO: 1128:
/ / SEQUENCE CHARACTERISTICS:
/ / LENGTH: 1798 base pairs
/ / TYPE: nucleic acid
/ / STRANDEDNESS: single
/ / TOPOLOGY: linear
/ / IMMEDIATE SOURCE:
/ / LIBRARY: GENBANK
/ / CLONE: G1399196
/ / US-09-016-434-1128

Query Match 6.5%; Score 71.8; DB 4; Length 1798;
Best Local Similarity 47.6%; Pred. No. 1.2e-13;
Matches 451; Conservative 0; Mismatches 442; Indels 54; Gaps 6;

OY 154 TTTCCTCACAAGAGGCTCCGATTCGGATCGGAGCGGCGATCGAACTCGAGCGCTG 213
DB 388 TTTTTCGCAAGAAATTCCTGAGACTGCGCAACATTCGAAAGAAATGATATCTC 447
OY 214 CCTTATGCGCTCTCGAGAAACCTGCGCTCTGAAAGTAAGATTTGGATTTGAGTCA 273
DB 448 CCGACCCATTTAGTAATACCTCTTCAAGTCAATTTGTTAAAGTGATATACAGC 507
OY 274 TTCAGGAGCATAGAGCGCTTCCTGATCA-----AGATATCTGCTGATGAGAAAG 327
DB 508 CTGATGAGATTTGGTGAATTCATAGAAAGCCAGATGACCAAGAAAGCATTTATCAAC 567
OY 328 TTCACACAGATGATCAAGGCTGTTAAAGTAAAGGCAACAAGTGTCCATGAGTGGCT 387
DB 568 TTTGTAGATACACTCATCAAGTTGGAATAGACCAATATGTATGCTCCATACAGCA 627
OY 388 CTGGGTGTGAACCACTGAAAGAAAGAAATGAAATCTTAC---GAAAGCTTGATGAGAT 444
DB 628 CAAAGAAATCATAGAGTATTAAGATGCTGTACAGTTGACCAAGTCAACATCAAAATCTT 687
OY 445 CATCAGTTCTTGATCGCTCTACTGTCTGTATAGGATCCGATGCTTATCGGCGAG 504
DB 688 CAATATTTCTTGATCGATTTTATCACTGACCGTATTTCTACTCGAGTGTGATGAACGAG 747
OY 505 CATGTTGAGTTGCATATCCAAACCCACACTTCAACA-----GTGGGTTACATACAC 558
DB 748 CACATTTCTTATATTAGTACCTCAGACAGAGAAACCAACCAATTGGAAGATGAT 807
OY 559 ACCAAGATGTCTCCATAGAGTGTGCAAGAAATGCTAGTAAAGTGAAGTGCAGTTGT 618

```



; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 377  
 ; LENGTH: 1422  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-949-016-377

Query Match 5.5%; Score 60.8; DB 4; Length 1422;  
 Best Local Similarity 48.4%; Pred. No. 7.1e-10;  
 Matches 457; Conservative 0; Mismatches 427; Indels 60; Gaps 8;

```

QY 154 TTCTTCACAGAGCTTCGATTCGATTCGAGAGCGGCGATGAACTGGAGACGGTG 213
DB 202 TTCCTCAGGCGAGAGCTGCTGTGGCGCTGGCCACATCATGAAGAATCACTGCTT 261
QY 214 CCTTATGAGCTCTCTGAGAAACCTGCCCTTTGAAGTAAGATTTGATGTGAGTCA 273
DB 262 CCGAGCCAGTGTCTGAGACACCCAGCTGAGCTGTGCAAGCTGTATGTCCAGAGC 321
QY 274 TTCAGGACATGAGACGTTTCTGAGA-----TCAGGATACGCTGATGAGAAAG 327
DB 322 CTCCTGAGACATGAGAGTTCTTGAACAAGATCCTGAGACCACTGACCCCTGAGCG 381
QY 328 TTCACACAGATGATCAAGCTGTAAAGTAAGGACAACAACGTTGTTCCATGATG 387
DB 382 TTCACGACGCTCTGTGATCAATCCGAAACCGGACAAACGATGATGATGCAATG 441
QY 388 CTGGGTGAGACGAGCTGAGAAAGAAAT---GAACTCTAGAAAGCTTATGATG 444
DB 442 CAGGCGCTCTGAGTACAGAGACCTACGCGAGTACCCCGTCTCAACCAAGACATC 501
QY 445 CATCATGTTCTGATCGCTTCTACTTGTCTGTATGAGGATCCGATATGCTTATCGG 504
DB 502 CAGTACTTCTGAGACCGCTTCTAAGCGGCACTTCCATTCGCAATGCTCATACAG 561
QY 505 -----CATGTTGAGTTGATATCAATCAACCCACCACTTACACAG-TGGTTA 555
DB 562 CACACCTCATCTTGTATGAGCAGACCAACCCAGCCATCCCAACATCGGAGCATC 621
QY 556 CACACCAAGATGTCTCTTATGAGTGAAGGAAAGAACTGTAAGTGAAGTGAAGTGA 615
DB 622 GACCCCAACTGCAACGCTCTGAGTGTGTAAAGTCTTCAAGCATGAGTAAAGTCT 681
QY 616 TGTTCAGAGATGAGTGTCTGCTCGAGATTAACATATATG-----658
DB 682 TGTGACAGATTAATGAGCTCTCACTGACCTGAGATCCAGGATCAATGAGCCAC 741
QY 659 -GGATCCAGTTTCACTTTTCCGATGTTCCGATCCGACCATTTGATGATGATGAT 717
DB 742 TCCAAGACCGGATTCACATGATGATGATGATGATGATGATGATGATGATGATG 801
QY 718 TTATGTAAGAACTCTCCGCTGCTGCTCAAGAGCGGTTTGTGATCTGATAGGTT 777
DB 802 CTCTTCAAGATGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 858
QY 778 CCAACCAATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 837
DB 859 CCAACCAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 918
QY 838 GAAGTGAAGTATACCGAAGACGGTCTCCAAATATTCCTAATCTTACCTTACAG 897
DB 919 CGAGTGGGGGTGTCTCTTGAAGAAATTTGAGGAGCTTTCAGCTATATGATTCACA 978
QY 898 GCAAGAAACCACTTGAAGAAATGAGACTTGGAAACCGCTGATGTTCCCTGACTAT 957
DB 979 GCAACCAACCCCTC-----AGCTGCAACCGGGGAGAGCGCGC-----TG 1017
QY 958 GCTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017
DB 1018 GCTGCTTGTGTTATGAGGCTCCCATTTCCGCTTCAAGCAAGTACTTCCAGGAGAG 1077
  
```

```

QY 1018 TTGCAATCATATTCATGAGAGATACGGAGCTGATCTTACTT 1061
DB 1078 CTGACGCTTCTTCCATGAGAAAGCTTTGGACCGAGCTGTAT 1121
  
```

# RESULT 15

```

US-09-949-016-17592
; Sequence 17592, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17592
; LENGTH: 72992
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17592
  
```

Query Match 5.5%; Score 60.8; DB 4; Length 72992;  
 Best Local Similarity 71.4%; Pred. No. 1.2e-08;  
 Matches 80; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```

QY 957 GCTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1016
DB 68219 GCTGATTTGTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1078
QY 1017 TTGCAATCATATTCATGAGAGATACGGAGCTGATCTTACTTGCACCTTG 1068
DB 68279 TGTGAACGTATTCATGAGAGAGTGGGATGATGATGATGATGATGATGAT 68330
  
```

Search completed: April 12, 2005, 05:29:23  
 Job time : 317 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 12, 2005, 05:36:03 ; Search time 1555 Seconds  
(without alignments)  
4302.408 Million cell updates/sec

Title: US-10-642-531-1

Perfect score: 1104  
1 atgcgcgtgaagaagcctag.....cgacggagccttcgcacga 1104

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 5615251 seqs, 3030001701 residues

Total number of hits satisfying chosen parameters: 11230502

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
21: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1104	100.0	1104	17	US-10-222-075-1
2	1104	100.0	1104	18	US-10-642-531-1
3	1095.6	99.2	1104	17	US-10-222-075-2
4	1095.6	99.2	1104	18	US-10-642-531-2
5	1062.4	96.2	1104	17	US-10-222-075-4
6	1062.4	96.2	1104	18	US-10-642-531-4
7	1062	96.2	1104	17	US-10-222-075-3
8	1062	96.2	1104	18	US-10-642-531-3
9	873.6	79.1	1457	14	US-10-202-428-1
10	665.8	60.3	1382	13	US-10-062-254-13
11	559.4	50.7	2028	17	US-10-424-599-109588

12	547.4	49.6	1481	13	US-10-062-254-9	Sequence 9, Appl1
13	547.4	49.6	2892	18	US-10-437-963-74020	Sequence 74020, A
14	536.6	48.6	1165	17	US-10-425-114-15131	Sequence 15131, A
15	527.8	47.8	1612	17	US-10-425-114-15532	Sequence 15532, A
16	527.8	47.8	1612	17	US-10-425-115-151648	Sequence 151648, A
17	524.8	47.5	1424	13	US-10-062-254-23	Sequence 23, Appl1
18	524.6	47.5	1467	17	US-10-425-114-32770	Sequence 32770, A
19	523.6	47.4	1876	18	US-10-767-701-14658	Sequence 14658, A
20	522.2	47.3	1388	17	US-10-425-114-34871	Sequence 34871, A
21	522.2	47.3	1469	18	US-10-425-115-94621	Sequence 94621, A
22	521.4	47.2	1467	18	US-10-425-115-151647	Sequence 151647, A
23	512	46.4	1910	18	US-10-437-963-17757	Sequence 17757, A
24	511.6	46.3	1407	13	US-10-062-254-3	Sequence 3, Appl1
25	511	46.3	1502	17	US-10-425-114-4107	Sequence 4107, Ap
26	441.6	40.0	1100	18	US-10-437-963-74022	Sequence 74022, A
27	419.6	38.0	1394	17	US-10-425-114-34658	Sequence 34658, A
28	367.8	33.3	1114	13	US-10-062-254-21	Sequence 21, Appl1
29	330.4	29.9	789	17	US-10-425-114-8176	Sequence 8176, Ap
30	328.2	29.7	591	18	US-10-021-332-9239	Sequence 9239, Ap
31	324.2	29.4	899	14	US-10-101-464A-965	Sequence 965, App
32	324.2	29.4	899	19	US-10-864-252-965	Sequence 2039, Ap
33	298.6	27.0	1378	18	US-10-739-930-2039	Sequence 110303, A
34	274.6	24.9	671	17	US-10-424-559-110303	Sequence 64166, A
35	274.4	24.9	2484	18	US-10-425-115-94166	Sequence 94622, A
36	266	24.1	2665	18	US-10-425-115-94622	Sequence 40472, A
37	263.8	23.9	790	17	US-10-424-559-40472	Sequence 109592, A
38	253	22.9	534	17	US-10-424-559-109592	Sequence 14799, A
39	253	22.9	734	17	US-10-425-114-14799	Sequence 7, Appl1
40	251.2	22.8	616	13	US-10-062-254-7	Sequence 1, Appl1
41	244.2	22.1	584	17	US-10-620-796-1	Sequence 11, Appl1
42	233.4	21.1	503	13	US-10-062-254-11	Sequence 110305, A
43	232	21.0	741	17	US-10-424-559-110305	Sequence 1, Appl1
44	219.6	19.9	665	13	US-10-062-254-1	Sequence 16529, A
45	218.4	19.8	704	17	US-10-425-114-16529	

## ALIGNMENTS

RESULT 1  
US-10-222-075-1  
; Sequence 1, Application US/10222075  
; Publication No. US20040033606A1  
; GENERAL INFORMATION:  
; APPLICANT: Zou, Jitao  
; TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE  
; FILE REFERENCE: 3015-544205  
; CURRENT APPLICATION NUMBER: US/10/222,075  
; CURRENT FILING DATE: 2002-08-16  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1104  
; TYPE: DNA  
; ORGANISM: Brassica napus  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: PDK cDNA from B. napus  
US-10-222-075-1

Query Match	100.0%	Score 1104;	DB 17;	Length 1104;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1104;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	ATGCGCGTGAAGAGCGCTAGCGAGATGTTTTCGAAGAGCTTATCGAGACCTTCACAGA	60	
DB	1	ATGCGCGTGAAGAGCGCTAGCGAGATGTTTTCGAAGAGCTTATCGAGACCTTCACAGA	60	
QY	61	TGGGATGATGAAGAGAGAGAGCGGCGTGAAGCTTCAGGATCATGATGAGTTCGATCACT	120	
DB	61	TGGGATGATGAAGAGAGAGAGCGGCGTGAAGCTTCAGGATCATGATGAGTTCGATCACT	120	
QY	121	CCCACTGAGAGAAACCTTCTGATCTCGGCGAGATTTCCTCACAGAGACTTCGATTCG	180	

```

Db      121 CCCATGAGAGAAACCTTCTGATCTCGGCGAGTTTCTTCAAGAGAGCTTCGATGCG
Qy      181 ATCCGAGAGGCGTCCGATCGAATCGAGAGCTGCTTATGAGCTCTGAGAAACCTGCC
Db      181 ATCCGAGAGGCGTCCGATCGAATCGAGAGCTGCTTATGAGCTCTGAGAAACCTGCC
Qy      241 GCTTGAAGGTAAAGATTGATGAGAGTCAATCAGGAGATAGAGCGCTTCCCTAG
Db      241 GCTTGAAGGTAAAGATTGATGAGAGTCAATCAGGAGATAGAGCGCTTCCCTAG
Qy      301 ATCAAGATATCTGCTGATGAGAAAGATTCAACAGATGATCAAGCTGTTAAAGTAAG
Db      301 ATCAAGATATCTGCTGATGAGAAAGATTCAACAGATGATCAAGCTGTTAAAGTAAG
Qy      361 CACACAAAGTGTCCCATGATGAGCTGCTGAGTGAACCGAGTCAAGAAAGATGAA
Db      361 CACACAAAGTGTCCCATGATGAGCTGCTGAGTGAACCGAGTCAAGAAAGATGAA
Qy      421 CTACGAAAAAGCTTGAATGAGATTGATGATGATGATGATGATGATGATGATGATGAT
Db      421 CTACGAAAAAGCTTGAATGAGATTGATGATGATGATGATGATGATGATGATGATGAT
Qy      481 GGGATCCCTATGCTTATGAGGAGAGATTTGATGATGATGATGATGATGATGATGATGAT
Db      481 GGGATCCCTATGCTTATGAGGAGAGATTTGATGATGATGATGATGATGATGATGATGAT
Qy      541 ACAGTGGGTTACATACACACCAAGATGCTCTTATGAGGAGGAGGAGGAGGAGGAGGAG
Db      541 ACAGTGGGTTACATACACACCAAGATGCTCTTATGAGGAGGAGGAGGAGGAGGAGGAG
Qy      601 GATCAAGAGTGAATTTGTTTCAAGAGATGATGATGATGATGATGATGATGATGATGAT
Db      601 GATCAAGAGTGAATTTGTTTCAAGAGATGATGATGATGATGATGATGATGATGATGAT
Qy      661 GATCAAGAGTGAATTTGTTTCAAGAGATGATGATGATGATGATGATGATGATGATGAT
Db      661 GATCAAGAGTGAATTTGTTTCAAGAGATGATGATGATGATGATGATGATGATGATGAT
Qy      721 GTCAAGAACTCTCTCGTGTCTCAAGAGCGGTTGTTGATGATGATGATGATGATGATGAT
Db      721 GTCAAGAACTCTCTCGTGTCTCAAGAGCGGTTGTTGATGATGATGATGATGATGATGAT
Qy      781 CCATCGGTATCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT
Db      781 CCATCGGTATCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT
Qy      841 GGTGAGGTATACCGAGAGCGGTTCTCTAAATATTAATTAATTAATTAATTAATTAAT
Db      841 GGTGAGGTATACCGAGAGCGGTTCTCTAAATATTAATTAATTAATTAATTAATTAAT
Qy      901 AGAAACCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      901 AGAAACCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      1021 CAGATCATATCATGAGAGATACGGAGCTGATGATGATGATGATGATGATGATGATGATGAT
Db      1021 CAGATCATATCATGAGAGATACGGAGCTGATGATGATGATGATGATGATGATGATGATGAT
Qy      1081 GACTGCAAGAGGCTTTGCCATGA 1104
Db      1081 GACTGCAAGAGGCTTTGCCATGA 1104

```

```

; APPLICANT: Marilia, Elizabeth-France
; APPLICANT: Zou, Jitao
; TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE
; CURRENT APPLICATION NUMBER: 3015-5442US
; CURRENT FILING DATE: 2003-08-15
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; OTHER INFORMATION: PDHK cDNA from B. napus
US-10-642-531-1

```

```

Query Match      100.0%; Score 1104; DB 18; Length 1104;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 ATGCGGTTAAGAGAGCTAGAGAGATGTTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAG
Db      1 ATGCGGTTAAGAGAGCTAGAGAGATGTTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAG
Qy      61 TGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      61 TGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      121 CCCATGAGAGAAACCTTCTGATCTCGGCGAGTTTCTTCAAGAGAGCTTCCATTCG
Db      121 CCCATGAGAGAAACCTTCTGATCTCGGCGAGTTTCTTCAAGAGAGCTTCCATTCG
Qy      181 ATGCGAGGCGTGCATGAACTCGAGAGCGTGCCTTATGAGCTCTCTGAGAAACCTGCC
Db      181 ATGCGAGGCGTGCATGAACTCGAGAGCGTGCCTTATGAGCTCTCTGAGAAACCTGCC
Qy      241 GCTTGAAGGTAAAGATTGATGAGAGTCAATCAGGAGATAGAGCGCTTCCCTAG
Db      241 GCTTGAAGGTAAAGATTGATGAGAGTCAATCAGGAGATAGAGCGCTTCCCTAG
Qy      301 ATCAAGATATCTGCTGATGAGAAAGATTCAACAGATGATCAAGCTGTTAAAGTAAG
Db      301 ATCAAGATATCTGCTGATGAGAAAGATTCAACAGATGATCAAGCTGTTAAAGTAAG
Qy      361 CACAAACAAGTGTCCCATGATGAGCTCTGAGTGAACCAAGTGAAGAAAGATGAA
Db      361 CACAAACAAGTGTCCCATGATGAGCTCTGAGTGAACCAAGTGAAGAAAGATGAA
Qy      421 CTACGAAAAAGCTTGAATGAGATTGATGATGATGATGATGATGATGATGATGATGATGAT
Db      421 CTACGAAAAAGCTTGAATGAGATTGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      481 GGGATCCCTATGCTTATGAGGAGAGATTTGATGATGATGATGATGATGATGATGATGAT
Db      481 GGGATCCCTATGCTTATGAGGAGAGATTTGATGATGATGATGATGATGATGATGATGAT
Qy      541 ACAGTGGGTTACATACACCAAGATGCTCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAG
Db      541 ACAGTGGGTTACATACACCAAGATGCTCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAG
Qy      601 GATCAAGAGTGAATTTGTTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT
Db      601 GATCAAGAGTGAATTTGTTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      661 GATCAAGAGTGAATTTGTTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT
Db      661 GATCAAGAGTGAATTTGTTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      721 GTCAAGAACTCTCTCGTGTCTCAAGAGCGGTTGTTGATGATGATGATGATGATGATGATGAT
Db      721 GTCAAGAACTCTCTCGTGTCTCAAGAGCGGTTGTTGATGATGATGATGATGATGATGATGAT

```

RESULT 2  
 US-10-642-531-1  
 ; Sequence 1, Application US/10642531  
 ; Publication No. US20040248302A1  
 ; GENERAL INFORMATION:



Db 721 GTCAGAACTCTCTCCGCTGCTCCAGAGCGTTTGTGACTGATAGGGTTGACCA 780  
Qy 781 CCAATCCGTAATGTTGCTGATGGAATGGAATGTTAATTAAGTCTGATGAA 840  
Db 781 CCAATCCGTAATGTTGCTGATGGAATGGAATGTTAATTAAGTCTGATGAA 840  
Qy 841 GGTGAGGTATACCGAAGAGGGTCTCCCTAAATATTCCTTACCTACAGACTGCA 900  
Db 841 GGTGAGGTATACCGAAGAGGGTCTCCCTAAATATTCCTTACCTACAGACTGCA 900  
Qy 901 AGAAACCCACTTGAAGAAAGATGAGTGGGAAACCGCTGATGTTCCCTGACTATGCT 960  
Db 901 AGAAACCCACTTGAAGAAAGATGAGTGGGAAACCGCTGATGTTCCCTGACTATGCT 960  
Qy 961 GGTATGTTATGTTGCTGCTGCTATGATGCTGATGCTGCTATGTTGGTGGAGATTG 1020  
Db 961 GGTATGTTATGTTGCTGCTGCTATGATGCTGATGCTGCTATGTTGGTGGAGATTG 1020  
Qy 1021 CAGATCATATCCATGGAAGATAGGGAAGTGAATGCTTACTTGCCTGCTTGTGA 1080  
Db 1021 CAGATCATATCCATGGAAGATAGGGAAGTGAATGCTTACTTGCCTGCTTGTGA 1080  
Qy 1081 GACTCGCAGAGAGCTTTGCCATGA 1104  
Db 1081 GACTCGCAGAGAGCTTTGCCATGA 1104

## RESULT 3

US-10-222-075-2  
; Sequence 2, Application US/10222075  
; Publication No. US20040033606A1  
; GENERAL INFORMATION:  
; APPLICANT: Zou, Jifao  
; TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE  
; FILE REFERENCE: 3015-5442US  
; CURRENT APPLICATION NUMBER: US/10/222,075  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 1104  
; TYPE: DNA  
; ORGANISM: Brassica rapa  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: PDHK cDNA from B. rapa  
US-10-222-075-2

Query Match 99.2%; Score 1095.6; DB 17; Length 1104;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1098; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 ATGGCGGTGAAGAGGCTTACGAGATGTTTTCGAAGCTTGTGATCGAGACGTTACAGA 60  
Db 1 ATGGCGGTGAAGAGGCTTACGAGATGTTTTCGAAGCTTGTGATCGAGACGTTACAGA 60  
Qy 61 TGGGAGATGATGAGAGAGAGGCGGTGAGGCTGACGATACATGATGAGTTCGATTCAC 120  
Db 61 TGGGAGATGATGAGAGAGAGGCGGTGAGGCTGACGATACATGATGAGTTCGATTCAC 120  
Qy 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAACAAGAGCTTCGATTCG 180  
Db 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAACAAGAGCTTCGATTCG 180  
Qy 181 ATGCGGAGGCTGCGATCGAATCGAGACGCTGCTTATGCGCTCTCTGAGAAACCTGCC 240  
Db 181 ATGCGGAGGCTGCGATCGAATCGAGACGCTGCTTATGCGCTCTCTGAGAAACCTGCC 240  
Qy 241 GTCTTGAAGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
Db 241 GTCTTGAAGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
Qy 301 ATCAAGATATCTGCTGATGAGAAAGATTCAACAAGATGATCAAGGCTGTTAAAGTAA 360

Db 301 ATCAAGATATCTGCTGATGAGAAAGATTCAACAAGATGATCAAGTGAATTAAGCTGTTAAAGTAA 360  
Qy 361 CACAAACAAGTGGTTCCTCATGATGCTGCTGGGTGTGAACCAAGCTGAAGAAAGAAATGAA 420  
Db 361 CACAAACAAGTGGTTCCTCATGATGCTGCTGGGTGTGAACCAAGCTGAAGAAAGAAATGAA 420  
Qy 421 CTCTAGAAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
Db 421 CTCTAGAAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
Qy 481 GGGATCGTATGCTTATCGGGCAGCATGTTGAGTGTGATATCAAAACCAACCACTTAC 540  
Db 481 GGGATCGTATGCTTATCGGGCAGCATGTTGAGTGTGATATCAAAACCAACCACTTAC 540  
Qy 541 ACAGTGGTATACATACACACCAAGATGCTCTCTATGAGAGTGGCAGGAATGCTAGTGA 600  
Db 541 ACAGTGGTATACATACACCAAGATGCTCTCTATGAGAGTGGCAGGAATGCTAGTGA 600  
Qy 601 GATGCAAGTGCATTTGTTTCAAGAGTATGTTGCTCGGAGATPAAATATATATG 660  
Db 601 GATGCAAGTGCATTTGTTTCAAGAGTATGTTGCTCGGAGATPAAATATATATG 660  
Qy 661 GATCAAGTTCATCTTTCCTGATGTTCCGACCACTTTCATGTTATGTTAGTGA 720  
Db 661 GATCAAGTTCATCTTTCCTGATGTTCCGACCACTTTCATGTTATGTTAGTGA 720  
Qy 721 GTCAGAACTCTCTCCGCTGCTCCAGAGCGGTTTGTGATCTGATAGGTTGACCA 780  
Db 721 GTCAGAACTCTCTCCGCTGCTCCAGAGCGGTTTGTGATCTGATAGGTTGACCA 780  
Qy 781 CCAATCCGTAATGTTGCTGATGGAATGGAATGTTCAATTAAGTCTGATGAA 840  
Db 781 CCAATCCGTAATGTTGCTGATGGAATGGAATGTTCAATTAAGTCTGATGAA 840  
Qy 841 GGTGAGGTATACCGAAGAGGGTCTCCCTAAATATTCCTTACTTACAGACTGCA 900  
Db 841 GGTGAGGTATACCGAAGAGGGTCTCCCTAAATATTCCTTACTTACAGACTGCA 900  
Qy 901 AGAAACCCACTTGAAGAAAGATGAGTGGGAAACCGCTGATGTTCCCTGACTATGCT 960  
Db 901 AGAAACCCACTTGAAGAAAGATGAGTGGGAAACCGCTGATGTTCCCTGACTATGCT 960  
Qy 961 GGTATGTTATGTTGCTGCTGCTATGATGCTGATGCTGCTATGTTGGTGGAGATTG 1020  
Db 961 GGTATGTTATGTTGCTGCTGCTATGATGCTGATGCTGCTATGTTGGTGGAGATTG 1020  
Qy 1021 CAGATCATATCCATGGAAGATAGGGAAGTGAATGCTTACTTGCCTGCTTGTGA 1080  
Db 1021 CAGATCATATCCATGGAAGATAGGGAAGTGAATGCTTACTTGCCTGCTTGTGA 1080  
Qy 1081 GACTCGCAGAGAGCTTTGCCATGA 1104  
Db 1081 GACTCGCAGAGAGCTTTGCCATGA 1104

## RESULT 4

US-10-642-531-2  
; Sequence 2, Application US/10642531  
; Publication No. US20040248302A1  
; GENERAL INFORMATION:  
; APPLICANT: Marillia, Elizabeth-France  
; APPLICANT: Zou, Jifao  
; TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE  
; FILE REFERENCE: 3015-5442US  
; CURRENT APPLICATION NUMBER: US/10/642,531  
; PRIOR APPLICATION NUMBER: 10/222,075  
; PRIOR FILING DATE: 2002-08-16  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2

LENGTH: 1104  
TYPE: DNA  
ORGANISM: Brassica rapa  
FEATURE:  
OTHER INFORMATION: PDHK cDNA from B. rapa  
US-10-642-531-2

Query Match 99.2%; Score 1095.6; DB 18; Length 1104;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1098; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```
OY 1 ATGGCGGTGAAGAGGCTTACGAGATGTTTGAAGAGCTTGATCGAGAGCTTCACAGA 60
DB 1 ATGGCGGTGAAGAGGCTTACGAGATGTTTGAAGAGCTTGATCGAGAGCTTCACAGA 60
OY 61 TGGGGATGATGAGAGAGCGGCGGTGAGCCCTGAGTCAATGATGAGATGCTGCTTCACT 120
DB 61 TGGGGATGATGAGAGAGCGGCGGTGAGCCCTGAGTCAATGATGAGATGCTGCTTCACT 120
OY 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGAGTTCTTCAAGAGAGCTTCGATTCGG 180
DB 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGAGTTCTTCAAGAGAGCTTCGATTCGG 180
OY 181 ATCGGAGGCGGTGAGTGAATGAACTGAGAGCGTGCCTTATGCGCTCTGAGAGAACTGCC 240
DB 181 ATCGGAGGCGGTGAGTGAATGAACTGAGAGCGTGCCTTATGCGCTCTGAGAGAACTGCC 240
OY 241 GTCTTGAAGTAAAGATTTGGTATGTGAGTCAATTCAGGAGCATGAGAGCGTTCCGAG 300
DB 241 GTCTTGAAGTAAAGATTTGGTATGTGAGTCAATTCAGGAGCATGAGAGCGTTCCGAG 300
OY 301 ATCAAGAGTACTGCTGATGAGAGAGTTCAACAGATGATCAAGGCTGTAAAGTAAAG 360
DB 301 ATCAAGAGTACTGCTGATGAGAGAGTTCAACAGATGATCAAGGCTGTAAAGTAAAG 360
OY 361 CACAACAACTGTGTTCCCATGATGAGTCTGGGTGAGAACAGCTGAGAGAGAGATGAAA 420
DB 361 CACAACAACTGTGTTCCCATGATGAGTCTGGGTGAGAACAGCTGAGAGAGAGATGAAA 420
OY 421 CTCTACGAAAAGCTTGAATGAGATTCATCAATTCCTGATGAGCTTCACTGCTCGATA 480
DB 421 CTCTACGAAAAGCTTGAATGAGATTCATCAATTCCTGATGAGCTTCACTGCTCGATA 480
OY 481 GGGATCGGTATGCTTATCGGCGAGAGTGAATGAGTGAATTCACCACTTCAC 540
DB 481 GGGATCGGTATGCTTATCGGCGAGAGTGAATGAGTGAATTCACCACTTCAC 540
OY 541 ACAAGTGGTAAATACACACCAAGATGCTCTATGAGAGTGGCAAGAAATGCTAGTGA 600
DB 541 ACAAGTGGTAAATACACACCAAGATGCTCTATGAGAGTGGCAAGAAATGCTAGTGA 600
OY 601 GATGCAAGTGCATTTGTTTCAAGAGTATGTTCTGCTCCGAGATTAACATATAATGCG 660
DB 601 GATGCAAGTGCATTTGTTTCAAGAGTATGTTCTGCTCCGAGATTAACATATAATGCG 660
OY 661 GATCAGAGTTCACTTTCCGTATGTTCCGAGCCCATTTGATGATTTATGATGATGATA 720
DB 661 GATCAGAGTTCACTTTCCGTATGTTCCGAGCCCATTTGATGATTTATGATGATGATA 720
OY 721 GTCAAGAACTCTCCGTGCTGTCAGAGCGGTTTGTGATCTGATAGAGGTTGACCA 780
DB 721 GTCAAGAACTCTCCGTGCTGTCAGAGCGGTTTGTGATCTGATAGAGGTTGACCA 780
OY 781 CCAATCGGTATCTGTTGCTGATGAGATGAGATGATTAAGATTAAGATTAAGATTAAG 840
DB 781 CCAATCGGTATCTGTTGCTGATGAGATGAGATGATTAAGATTAAGATTAAGATTAAG 840
OY 841 GGTGAGGTATACCGAGAGCGGCTTCCTTAAATATTAATTAATTAATTAATTAATTA 900
DB 841 GGTGAGGTATACCGAGAGCGGCTTCCTTAAATATTAATTAATTAATTAATTAATTA 900
OY 901 AGAAACCCACTGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 960
DB 901 AGAAACCCACTGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 960
```

```
DB 901 AGAAACCCACTGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 960
OY 961 GGTATGATTAATGATGCTGCTTATGATGCTGCTTATGATGCTGCTTATGATGCTGCTT 1020
DB 961 GGTATGATTAATGATGCTGCTTATGATGCTGCTTATGATGCTGCTTATGATGCTGCTT 1020
OY 1021 CAGATCAATTCATGAGAGAGATGAGAGATGAGATGAGATGAGATGAGATGAGATGAG 1080
DB 1021 CAGATCAATTCATGAGAGAGATGAGAGATGAGATGAGATGAGATGAGATGAGATGAG 1080
OY 1081 GACTCGAGAGAGCTTGGCATGA 1104
DB 1081 GACTCGAGAGAGCTTGGCATGA 1104
```

RESULT 5  
US-10-222-075-4  
Sequence 4, Application US/10222075  
Publication No. US20040033606A1  
GENERAL INFORMATION:  
APPLICANT: Zou, Jitao  
TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE  
FILE REFERENCE: 3015-5442US  
CURRENT APPLICATION NUMBER: US/10/222,075  
CURRENT FILING DATE: 2002-08-16  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 4  
LENGTH: 1104  
TYPE: DNA  
ORGANISM: Brassica carinata  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: PDHK cDNA from B. carinata  
US-10-222-075-4

Query Match 96.2%; Score 1062.4; DB 17; Length 1104;  
Best Local Similarity 97.6%; Pred. No. 0;  
Matches 1078; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```
OY 1 ATGGCGGTGAAGAGGCTTACGAGATGTTTGAAGAGCTTGATCGAGAGCTTCACAGA 60
DB 1 ATGGCGGTGAAGAGGCTTACGAGATGTTTGAAGAGCTTGATCGAGAGCTTCACAGA 60
OY 61 TGGGGATGATGAGAGAGCGGCGGTGAGCCCTGAGTCAATGATGAGATGCTGCTTCACT 120
DB 61 TGGGGATGATGAGAGAGCGGCGGTGAGCCCTGAGTCAATGATGAGATGCTGCTTCACT 120
OY 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGAGTTCTTCAAGAGAGCTTCGATTCGG 180
DB 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGAGTTCTTCAAGAGAGCTTCGATTCGG 180
OY 181 ATCGGAGGCGGTGAGTGAATGAACTGAGAGCGTGCCTTATGCGCTCTGAGAGAACTGCC 240
DB 181 ATCGGAGGCGGTGAGTGAATGAACTGAGAGCGTGCCTTATGCGCTCTGAGAGAACTGCC 240
OY 241 GTCTTGAAGTAAAGATTTGGTATGTGAGTCAATTCAGGAGCATGAGAGCGTTCCGAG 300
DB 241 GTCTTGAAGTAAAGATTTGGTATGTGAGTCAATTCAGGAGCATGAGAGCGTTCCGAG 300
OY 301 ATCAAGAGTACTGCTGATGAGAGAGTTCAACAGATGATCAAGGCTGTAAAGTAAAG 360
DB 301 ATCAAGAGTACTGCTGATGAGAGAGTTCAACAGATGATCAAGGCTGTAAAGTAAAG 360
OY 361 CACAACAACTGTGTTCCCATGATGAGTCTGGGTGAGAACAGCTGAGAGAGAGATGAAA 420
DB 361 CACAACAACTGTGTTCCCATGATGAGTCTGGGTGAGAACAGCTGAGAGAGAGATGAAA 420
OY 421 CTCTACGAAAAGCTTGAATGAGATTCATCAATTCCTGATGAGCTTCACTGCTCGATA 480
DB 421 CTCTACGAAAAGCTTGAATGAGATTCATCAATTCCTGATGAGCTTCACTGCTCGATA 480
OY 481 GGGATCGGTATGCTTATCGGCGAGAGTGAATGAGTGAATTCACCACTTCAC 540
DB 481 GGGATCGGTATGCTTATCGGCGAGAGTGAATGAGTGAATTCACCACTTCAC 540
```

Db 481 GGGATCGGTATGCTTATCGGGCAGACATGTTGATGATATCCAAACCCACACTTAC 540  
Qy 541 ACAGTGGTATACATACACCAAGATGCTCTATGAGGTCGACAAAGAAATGCTAGTGA 600  
Db 541 ACTGTGGTATACATACACCAAGATGCTCTATGAGGTCGACAAAGAAATGCTAGTGA 600  
Qy 601 GATGCAAGTGCATGTTGTTCCAGAGATGATGTTCTGCTCCGAGATTAACAATATATG 660  
Db 601 GATGCAAGTGCATGTTGTTCCAGAGATGATGTTCTGCTCCGAGATTAACAATATATG 660  
Qy 661 GATCAGAGTTCACTTTCCGATGTTCCGACCCATTTGACCTTATGTTGATGATTA 720  
Db 661 GATCAGAGTTCACTTTCCGATGTTCCGACCCATTTGACCTTATGTTGATGATTA 720  
Qy 721 GTCAGAACTCTCTCCGTCGTCGTCAGAGGCTTTGTTGATCTGATAGGTTGCACCA 780  
Db 721 GTCAGAACTCTCTACGTCGTCGTCAGAGGCTTTGTTGATCTGATAGGTTGCACCA 780  
Qy 781 CCAATCCGATATCTGTTGTCGATGATGATGATGATGATGATGATGATGATGATG 840  
Db 781 CCAATCCGATATCTGTTGTCGATGATGATGATGATGATGATGATGATGATGATG 840  
Qy 841 GGTGAGATATACCGAGAGCGGCTCTCCATAATATCTTACCTTACAGCATGCA 900  
Db 841 GGTGAGATATACCGAGAGCGGCTCTCCATAATATCTTACCTTACAGCATGCA 900  
Qy 901 AGAAACCCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 960  
Db 901 AGAAACCCGCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 960  
Qy 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
Db 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
Qy 1021 CAGATCATATCATGAGAGATACGGAGCTGATGATGATGATGATGATGATGATGAT 1080  
Db 1021 CAGATCATATCATGAGAGATACGGAGCTGATGATGATGATGATGATGATGATGAT 1080  
Qy 1081 GACTCGAGAGAGCTTTGCGATGA 1104  
Db 1081 GACTCGAGAGAGCTTTGCGATGA 1104

RESULT 6  
US-10-642-531-4  
Sequence 4, Application US/10642531  
Publication No. US20040248302A1  
GENERAL INFORMATION:  
APPLICANT: Marifilia, Elizabeth-France  
APPLICANT: Zou, Jileao  
APPLICANT: Taylor, David C.  
TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE  
FILE REFERENCE: 3015-5442US  
CURRENT APPLICATION NUMBER: US/10/642,531  
PRIOR FILING DATE: 2003-08-15  
PRIOR APPLICATION NUMBER: 10/222,075  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 4  
LENGTH: 1104  
TYPE: DNA  
ORGANISM: *Brassica carinata*  
FEATURE:  
OTHER INFORMATION: PDHK cDNA from *B. carinata*  
US-10-642-531-4

Query Match 96.2%; Score 1062.4; DB 18; Length 1104;  
Best Local Similarity 97.6%; Pred. No. 0;  
Matches 1078; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 ATGCGCGTGAAGAGGCTGACGAGATGTTTTCAGAGAGCTTGATCGAGAGCTTCAAGA 60

Db 1 ATGCGCGTGAAGAGGCTGACGAGATGTTTTCAGAGAGCTTGATCGAGAGCTTCAAGA 60  
Qy 61 TGGGATGATCAATGAAGCAGAGGCGCTGAGCTCAGAGTACATGATGAGATTCAGTCACT 120  
Db 61 TGGGATGATCAATGAAGCAGAGGCGCTGAGCTCAGAGTACATGATGAGATTCAGTCACT 120  
Qy 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAGAGCTTCGATTCGG 180  
Db 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAGAGCTTCGATTCGG 180  
Qy 181 ATGCGAGAGGCTGCGATCGAACTCGAGAGCTGCTTATGAGCTCTTGAAGAACTGCGC 240  
Db 181 ATGCGAGAGGCTGCGATCGAACTCGAGAGCTGCTTATGAGCTCTTGAAGAACTGCGC 240  
Qy 241 GTCCTGAAGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
Db 241 GTCCTGAAGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
Qy 301 ATCAAGATATCTGCTGATGAGAGAGTTCACACAGATGATCAAGGCTGTTAAAGTAAG 360  
Db 301 ATCAAGATATCTGCTGATGAGAGAGTTCACACAGATGATCAAGGCTGTTAAAGTAAG 360  
Qy 361 CACAGAAAGTGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420  
Db 361 CACAGAAAGTGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420  
Qy 421 CTCTAGAGAAAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
Db 421 CTCTAGAGAAAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
Qy 481 GGGATCCGATGCTTATGCGGCGAGAGTGTGATGATGATGATGATGATGATGATGATG 540  
Db 481 GGGATCCGATGCTTATGCGGCGAGAGTGTGATGATGATGATGATGATGATGATGATG 540  
Qy 541 ACAGTGGTATACATACACCAAGATGCTCTATGAGGTCGACAAAGAAATGCTAGTGA 600  
Db 541 ACAGTGGTATACATACACCAAGATGCTCTATGAGGTCGACAAAGAAATGCTAGTGA 600  
Qy 601 GATGCAAGTGCATGTTGTTCCAGAGATGATGATGATGATGATGATGATGATGATGATG 660  
Db 601 GATGCAAGTGCATGTTGTTCCAGAGATGATGATGATGATGATGATGATGATGATGATG 660  
Qy 661 GATCAGAGTTCACTTTCCGATGTTCCGACCCATTTGACCTTATGTTGATGATGATG 720  
Db 661 GATCAGAGTTCACTTTCCGATGTTCCGACCCATTTGACCTTATGTTGATGATGATG 720  
Qy 721 GTCAGAACTCTCTCCGTCGTCGTCAGAGGCTTTGTTGATCTGATAGGTTGCACCA 780  
Db 721 GTCAGAACTCTCTACGTCGTCGTCAGAGGCTTTGTTGATCTGATAGGTTGCACCA 780  
Qy 781 CCAATCCGATATCTGTTGTCGATGATGATGATGATGATGATGATGATGATGATGATG 840  
Db 781 CCAATCCGATATCTGTTGTCGATGATGATGATGATGATGATGATGATGATGATGATG 840  
Qy 841 GGTGAGATATACCGAGAGCGGCTCCGTAATAATATCTTACCTTACAGCATGCA 900  
Db 841 GGTGAGATATACCGAGAGCGGCTCCGTAATAATATCTTACCTTACAGCATGCA 900  
Qy 901 AGAAACCCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 960  
Db 901 AGAAACCCGCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 960  
Qy 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
Db 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
Qy 1021 CAGATCATATCATGAGAGATACGGAGCTGATGATGATGATGATGATGATGATGATG 1080  
Db 1021 CAGATCATATCATGAGAGATACGGAGCTGATGATGATGATGATGATGATGATGATG 1080  
Qy 1081 GACTCGAGAGAGCTTTGCGATGA 1104  
Db 1081 GACTCGAGAGAGCTTTGCGATGA 1104

Db 1081 GACTCGCAGAGCCTTGCATGA 1104

RESULT 7

US-10-222-075-3

Sequence 3, Application US/10222075

Publication No. US20040033606A1

GENERAL INFORMATION:

APPLICANT: Zou, Jitao

TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE

FILE REFERENCE: 3015-5442US

CURRENT APPLICATION NUMBER: US/10/222,075

CURRENT FILING DATE: 2002-08-16

NUMBER OF SEQ ID NOS: 15

SOFTWARE: PatentIn version 3.1

SEQ ID NO 3

LENGTH: 1104

TYPE: DNA

ORGANISM: Brassica oleracea

FEATURE:

NAME/KEY: misc.feature

OTHER INFORMATION: PDHK cDNA from B. oleracea

US-10-222-075-3

Query Match 96.2%; Score 1062; DB 17; Length 1104;

Best Local Similarity 97.0%; Pred. No. 0;

Matches 1071; Conservative 9; Mismatches 24; Indels 0; Gaps 0;

Db 1 ATGGCGGTGAAGAGGCTAGCAAGATGTTTTCGAAGAGCTTGATCGAGAGCTTCACAGA 60

1 ATGGCGGTGAAGAGGCTAGCAAGATGTTTTCGAAGAGCTTGATCGAGAGCTTCACAGA 60

61 TGGGATGATGAAAGCAGAGCGGGCTGAGCCTCAGTACATGATGAGATTGCTCCACT 120

1 TGGGATGATGAAAGCAGAGCGGGCTGAGCCTCAGTACATGATGAGATTGCTCCACT 120

61 TGGGATGATGAAAGCAGAGCGGGCTGAGCCTCAGTACATGATGAGATTGCTCCACT 120

121 CCCACTGAGAGAAACCTTCGATCTCGGCGCAGTTTCTTCAAGAGGCTTCGATTCGG 180

121 CCCACTGAGAGAAACCTTCGATCTCGGCGCAGTTTCTTCAAGAGGCTTCGATTCGG 180

121 CCCACTGAGAGAAACCTTCGATCTCGGCGCAGTTTCTTCAAGAGGCTTCGATTCGG 180

181 ATCGCAGAGCGGTGCGATGAACTCGAGACGCTGCTTATGAGCCTCTCTGAGAAACCTGCC 240

181 ATCGCAGAGCGGTGCGATGAACTCGAGACGCTGCTTATGAGCCTCTCTGAGAAACCTGCC 240

181 ATCGCAGAGCGGTGCGATGAACTCGAGACGCTGCTTATGAGCCTCTCTGAGAAACCTGCC 240

241 GCTTGAAGTGAAGATGGTATGATGAGTCAATTCAGGACATGAGAGCGTTTCTGAG 300

241 GCTTGAAGTGAAGATGGTATGATGAGTCAATTCAGGACATGAGAGCGTTTCTGAG 300

241 GCTTGAAGTGAAGATGGTATGATGAGTCAATTCAGGACATGAGAGCGTTTCTGAG 300

301 ATCAAGATATCTGCTGATGAGAAAGATTCACACAGATGATCAAGGCTGTTAAGTAAAG 360

301 ATCAAGATATCTGCTGATGAGAAAGATTCACACAGATGATCAAGGCTGTTAAGTAAAG 360

301 ATCAAGATATCTGCTGATGAGAAAGATTCACACAGATGATCAAGGCTGTTAAGTAAAG 360

361 CACAAACAAGTGGTCCCATGATGCTCTGGGTGTGAACCAAGTGAAGAAAGATGAA 420

361 CACAAACAAGTGGTCCCATGATGCTCTGGGTGTGAACCAAGTGAAGAAAGATGAA 420

361 CACAAACAAGTGGTCCCATGATGCTCTGGGTGTGAACCAAGTGAAGAAAGATGAA 420

421 CTCTACGAAAGAGCTTGATGAGATTCATCAGTTCTTGATGCTTCACTTCTGATTA 480

421 CTCTACGAAAGAGCTTGATGAGATTCATCAGTTCTTGATGCTTCACTTCTGATTA 480

421 CTCTACGAAAGAGCTTGATGAGATTCATCAGTTCTTGATGCTTCACTTCTGATTA 480

481 GGGATCCGTATGCTTATCGGGCAGCATGTTGAGTTGCATATCAATCAACCAACCACTTCA 540

481 GGGATCCGTATGCTTATCGGGCAGCATGTTGAGTTGCATATCAATCAACCAACCACTTCA 540

481 GGGATCCGTATGCTTATCGGGCAGCATGTTGAGTTGCATATCAATCAACCAACCACTTCA 540

541 ACAGTGGGTATCATCAACCAAGATGCTCTCTATGAGAGTGGCAAGAAATCTTATGTA 600

541 ACAGTGGGTATCATCAACCAAGATGCTCTCTATGAGAGTGGCAAGAAATCTTATGTA 600

541 ACAGTGGGTATCATCAACCAAGATGCTCTCTATGAGAGTGGCAAGAAATCTTATGTA 600

601 GATGCAAGTGCATTTGTTTCAGAGATGATGTTGCTCGGAGATTAACATATATGTC 660

601 GATGCAAGTGCATTTGTTTCAGAGATGATGTTGCTCGGAGATTAACATATATGTC 660

601 GATGCAAGTGCATTTGTTTCAGAGATGATGTTGCTCGGAGATTAACATATATGTC 660

661 GATCACAAGTTTCACTTTCCGTATGTTCCGACCACTTTGCAATTAATGATGATGATTA 720

661 GATCACAAGTTTCACTTTCCGTATGTTCCGACCACTTTGCAATTAATGATGATGATTA 720

Db 661 GATCACAAGTTTCACTTTCCGTATGTTCCGACCACTTTGCAATTAATGATGATGATTA 720

661 GATCACAAGTTTCACTTTCCGTATGTTCCGACCACTTTGCAATTAATGATGATGATTA 720

721 GTCAAGAACTCTCTCGGTGCTGCAAGAGCGGTTTGACCTCTATAGGTTGACCA 780

721 GTCAAGAACTCTCTCGGTGCTGCAAGAGCGGTTTGACCTCTATAGGTTGACCA 780

721 GTCAAGAACTCTCTCGGTGCTGCAAGAGCGGTTTGACCTCTATAGGTTGACCA 780

781 CCAATCCGTATCATTTGCTGATGAGATTCAGATGTTACATTAAGAGTTCAGATGAA 840

781 CCAATCCGTATCATTTGCTGATGAGATTCAGATGTTACATTAAGAGTTCAGATGAA 840

781 CCAATCCGTATCATTTGCTGATGAGATTCAGATGTTACATTAAGAGTTCAGATGAA 840

841 GGTGAGGTATACCGAAGAGCGGTTCCCTAAATATTAATTAATTAATTAATTAATTA 900

841 GGTGAGGTATACCGAAGAGCGGTTCCCTAAATATTAATTAATTAATTAATTAATTA 900

841 GGTGAGGTATACCGAAGAGCGGTTCCCTAAATATTAATTAATTAATTAATTAATTA 900

901 AGAACCACCTGGAAGAGATGAGACTTGGGAAACCGCTGATGTTCCCTGACTATGGCT 960

901 AGAACCACCTGGAAGAGATGAGACTTGGGAAACCGCTGATGTTCCCTGACTATGGCT 960

901 AGAACCACCTGGAAGAGATGAGACTTGGGAAACCGCTGATGTTCCCTGACTATGGCT 960

961 GGTATGATGATGATGCTGCTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020

961 GGTATGATGATGATGCTGCTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020

961 GGTATGATGATGATGCTGCTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020

1021 CAGATCATATCCATGGAAGAGATGAGACTGATGCTTACTTGCATGCTGCTGCTGCTG 1080

1021 CAGATCATATCCATGGAAGAGATGAGACTGATGCTTACTTGCATGCTGCTGCTGCTG 1080

1021 CAGATCATATCCATGGAAGAGATGAGACTGATGCTTACTTGCATGCTGCTGCTGCTG 1080

1081 GACTCGCAGAGCCTTTGCCATGA 1104

1081 GACTCGCAGAGCCTTTGCCATGA 1104

RESULT 8

US-10-642-531-3

Sequence 3, Application US/10642531

Publication No. US20040248302A1

GENERAL INFORMATION:

APPLICANT: Marillia, Elizabeth-France

APPLICANT: Zou, Jitao

TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE

FILE REFERENCE: 3015-5442US

CURRENT APPLICATION NUMBER: US/10/642,531

PRIOR FILING DATE: 2003-08-15

PRIOR APPLICATION NUMBER: 10/222,075

NUMBER OF SEQ ID NOS: 19

SOFTWARE: PatentIn version 3.1

SEQ ID NO 3

LENGTH: 1104

TYPE: DNA

ORGANISM: Brassica oleracea

FEATURE:

OTHER INFORMATION: PDHK cDNA from B. oleracea

US-10-642-531-3

Query Match 96.2%; Score 1062; DB 18; Length 1104;

Best Local Similarity 97.0%; Pred. No. 0;

Matches 1071; Conservative 9; Mismatches 24; Indels 0; Gaps 0;

Db 1 ATGGCGGTGAAGAGGCTAGCAAGATGTTTTCGAAGAGCTTGATCGAGAGCTTCACAGA 60

1 ATGGCGGTGAAGAGGCTAGCAAGATGTTTTCGAAGAGCTTGATCGAGAGCTTCACAGA 60

1 ATGGCGGTGAAGAGGCTAGCAAGATGTTTTCGAAGAGCTTGATCGAGAGCTTCACAGA 60

61 TGGGATGATGAAAGCAGAGCGGGCTGAGCCTCAGTACATGATGAGATTGCTCCACT 120

61 TGGGATGATGAAAGCAGAGCGGGCTGAGCCTCAGTACATGATGAGATTGCTCCACT 120

61 TGGGATGATGAAAGCAGAGCGGGCTGAGCCTCAGTACATGATGAGATTGCTCCACT 120

121 CCCACTGAGAGAAACCTTCGATCTCGGCGCAGTTTCTTCAAGAGGCTTCGATTCGG 180

121 CCCACTGAGAGAAACCTTCGATCTCGGCGCAGTTTCTTCAAGAGGCTTCGATTCGG 180

121 CCCACTGAGAGAAACCTTCGATCTCGGCGCAGTTTCTTCAAGAGGCTTCGATTCGG 180

181 ATCGCAGAGCGGTGCGATGAACTCGAGACGCTGCTTATGAGCCTCTCTGAGAAACCTGCC 240

181 ATCGCAGAGCGGTGCGATGAACTCGAGACGCTGCTTATGAGCCTCTCTGAGAAACCTGCC 240

```

Db 181 ATGCGAGGGGTCGATCGAACTCGAGACCGCTGCTTATGGCTCTCTGAGAAACCTGCC 240
Qy 241 GTCTTGAAGGTAGAGATTGCTATGTGGATCATTTGAGGACATGAGAGGCTTCCCTGAG 300
Db 241 GTCTTGAAGGTAGAGATTGCTATGTGGATCATTTGAGGACATGAGAGGCTTCCCTGAG 300
Qy 301 ATCAAGGATCTGCTGATGAGAAAGAGTTCACACAGATGATCAAGGCTGTTAAAGTAGG 360
Db 301 ATCAAGGATCTGCTGATGAGAAAGAGTTCACACAGATGATCAAGGCTGTTAAAGTAGG 360
Qy 361 CACAACAACGTGCTTCCATGATGCTCTGAGTGTGAACCACTGAAGAAAGAAATGAAA 420
Db 361 CACAACAACGTGCTTCCATGATGCTCTGAGTGTGAACCACTGAAGAAAGAAATGAAA 420
Qy 421 CTCTACGAAAAGCTGATGAGATTCATCACTTCTTGTATGCTTCTTCTCTGCTGATA 480
Db 421 CTCTACGAAAAGCTGATGAGATTCATCACTTCTTGTATGCTTCTTCTCTGCTGATA 480
Qy 481 GGGATCCGTATGCTTATCGGGCAGCATGTGATGATCAATCCAAACCCACACTTCCAC 540
Db 481 GGGATCCGTATGCTTATCGGGCAGCATGTGATGATCAATCCAAACCCACACTTCCAC 540
Qy 541 ACAGTGGTTCATATACACACCAAGATGCTCTATGAGAGGTGCAAGAAATGCTATGAA 600
Db 541 ACTGTGGTTCATATACACCAAGATGCTCTATGAGAGGTGCAAGAAATGCTATGAA 600
Qy 601 GATGCAAGGTGATTTGTTTTCAGAGAGTATGCTTCTGCTCCGAGATTAACATATATGCG 660
Db 601 GATGCAAGGTGATTTGTTTTCAGAGAGTATGCTTCTGCTCCGAGATTAACATATATGCG 660
Qy 661 GATCCAAGTTTCACTTTTCCGTATGTTCCGACCAATTTGATCTTATGATGATGATA 720
Db 661 GATCCAAGTTTCACTTTTCCGTATGTTCCGACCAATTTGATCTTATGATGATGATA 720
Qy 721 GTCAAGAACTCTCTCCGTGCTTCAGAGCGGTTGTTGATCTTGATGAGGTTGACCA 780
Db 721 GTCAAGAACTCTCTCCGTGCTTCAGAGCGGTTGTTGATCTTGATGAGGTTGACCA 780
Qy 781 CCAATCCGTATCATTTGTTGCTGATGAAATGAGAAATCTTCAATTAAGGCTCAGATGAA 840
Db 781 CCAATCCGTATCATTTGTTGCTGATGAAATGAGAAATCTTCAATTAAGGCTCAGATGAA 840
Qy 841 GGTGAGGTATATACGAGAGCGGCTCTCTAAATATTTCACTTACCTCAGACATGCA 900
Db 841 GGTGAGGTATATACGAGAGCGGCTCTCTAAATATTTCACTTACCTCAGACATGCA 900
Qy 901 AGAAACCCACTTGAAGAAATGATGAACTTGGGAAACCGCTGATGTTCCCTGACTATGCT 960
Db 901 AGAAACCCACTTGAAGAAATGATGAACTTGGGAAACCGCTGATGTTCCCTGACTATGCT 960
Qy 961 GGTATAGGTATAGTCTGCTATTAATGCTGCTTGTATGCTGCTATTTTGTGAGATGTTG 1020
Db 961 GGTATAGGTATAGTCTGCTATTAATGCTGCTTGTATGCTGCTATTTTGTGAGATGTTG 1020
Qy 1021 CAGATCATATCATGAGAGATACGGGATGATGCTTACTTGTGCACTTGTCTGCTTGA 1080
Db 1021 CAGATCATATCATGAGAGATACGGGATGATGCTTACTTGTGCACTTGTCTGCTTGA 1080
Qy 1081 GACTCGCAGAGAGCTTTTGCATGA 1104
Db 1081 GACTCGCAGAGAGCTTTTGCATGA 1104

```

```

RESULT 9
US-10-202-428-1
; Sequence 1, Application US/10202428
; Publication No. US20030084472A1
; GENERAL INFORMATION:
; APPLICANT: Zou, Jitao
; APPLICANT: Taylor, David
; TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE
; FILE REFERENCE: 40942-1

```

```

; CURRENT APPLICATION NUMBER: US/10/202,428
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/09/355,912
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/CA98/00096
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/038,815
; PRIOR FILING DATE: 1997-02-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1457
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-202-428-1

Query Match 79.1%; Score 873.6; DB 14; Length 1457;
Best Local Similarity 87.6%; Pred. No. 2,9e-273;
Matches 967; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

Qy 1 ATGCGGTGAGAAAGCTAGCGATGTTTTCGAAAGCTGATCGAGAGCTTCAAGA 60
Db 104 ATGCGGTGAGAAAGCTAGCGATGTTTTCGAAAGCTGATCGAGAGCTTCAAGA 163
Qy 61 TGGGATGCAATGAGAGACAGCGGCTGAGGCTCAGGTACATGATGATGATGATGAT 120
Db 164 TGGGATGCAATGAGAGACAGCGGCTGAGGCTCAGGTACATGATGATGATGATGATGAT 223
Qy 121 CCCACTGAGAGAAACCTTCTGATCTGCGGCAAGTTTCTTCAAGAGCTTCCGATTCCG 180
Db 224 CTTACTGAGAGAAATCTTCTGATCTGCGGCAAGTTTCTTCAAGAGCTTCCGATTCCG 283
Qy 181 ATGCGAGGGGTCGATCGAACTCGAGAGCTGCTGATGATGATGATGATGATGATGATGAT 240
Db 284 GTGCGAGAGAGAGATCGAACTCGAGAGCTGCTGATGATGATGATGATGATGATGATGATGAT 343
Qy 344 GTTTTGAAGGTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 403
Qy 404 ATCAAGATATCTGCTGATGAGAAAGCTTCAACAGATGATCAAGGCTGTTAAAGTAGG 360
Db 404 ATCAAGATATCTGCTGATGAGAAAGCTTCAACAGATGATCAAGGCTGTTAAAGTAGG 463
Qy 464 ATCAAGATATCTGCTGATGAGAAAGCTTCAACAGATGATCAAGGCTGTTAAAGTAGG 420
Db 522 ATCAAGATATCTGCTGATGAGAAAGCTTCAACAGATGATCAAGGCTGTTAAAGTAGG 522
Qy 522 ATCAAGATATCTGCTGATGAGAAAGCTTCAACAGATGATCAAGGCTGTTAAAGTAGG 580
Db 580 ATCAAGATATCTGCTGATGAGAAAGCTTCAACAGATGATCAAGGCTGTTAAAGTAGG 640
Qy 640 ATCAAGATATCTGCTGATGAGAAAGCTTCAACAGATGATCAAGGCTGTTAAAGTAGG 600
Db 641 ATCAAGATATCTGCTGATGAGAAAGCTTCAACAGATGATCAAGGCTGTTAAAGTAGG 700
Qy 700 ATCAAGATATCTGCTGATGAGAAAGCTTCAACAGATGATCAAGGCTGTTAAAGTAGG 660
Db 701 ATCAAGATATCTGCTGATGAGAAAGCTTCAACAGATGATCAAGGCTGTTAAAGTAGG 760
Qy 761 ATCAAGATATCTGCTGATGAGAAAGCTTCAACAGATGATCAAGGCTGTTAAAGTAGG 720
Db 761 ATCAAGATATCTGCTGATGAGAAAGCTTCAACAGATGATCAAGGCTGTTAAAGTAGG 820
Qy 821 ATCAAGATATCTGCTGATGAGAAAGCTTCAACAGATGATCAAGGCTGTTAAAGTAGG 780
Db 821 ATCAAGATATCTGCTGATGAGAAAGCTTCAACAGATGATCAAGGCTGTTAAAGTAGG 880
Qy 881 ATCAAGATATCTGCTGATGAGAAAGCTTCAACAGATGATCAAGGCTGTTAAAGTAGG 840

```





Db 1048 ACAATGGCTGATATGATATGCTCTTATTAGTCTCTATATAGTCTCGATTATTTTGA 1107  
 QY 1012 GGAGATTGGAGATCATATCCATGATGAAGATAGGGAGTGAATGCTTACTTGCATTGCT 1071  
 Db 1108 GGTGATCTTCAAAATTAATCTTATGAGAGATATGGCACTGATGATATATCTCCATTGTCT 1167  
 QY 1072 CGTCTTGAGACTCGCAGAGAGCCTTGGCATTGA 1104  
 Db 1168 CGTTGGAGATTCAACAAGAACCTTGGCTTGA 1200  
  
 RESULT 11  
 US-10-424-599-109588  
 ; Sequence 109588, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53223) B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 109588  
 ; LENGTH: 2028  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_69973C.1  
 US-10-424-599-109588  
  
 Query Match 50.7%; Score 559.4; DB 17; Length 2028;  
 Best Local Similarity 70.9%; Pred. No. 9,7e-171;  
 Matches 789; Conservative 0; Mismatches 311; Indels 13; Gaps 3;

QY 532 CCACTTCAACAGTGGATTACATACACCAAGATGCTCTATNGAGTGGCAGAGAT 591  
 Db 949 TCTCCCAAGTGTGGCTATATACATACAAAATGCTCTGTGAGATGACAGAGAT 1008  
 QY 592 GCTAGTGAAGATCAAGTGTGATTTTTCAGAGATAGTGTCTGCTCCGAGATTAAC 651  
 Db 1009 GCCAGTGAAGATGACAGTGTATATGTGTGCGGAATATGAGATGCGCCCTGATGTCAAT 1068  
 QY 652 ATATATGGGATTCAGATTTTCCATTTCCGTATGTTCCGACCATTTGATCTTATGCTG 711  
 Db 1069 ATTATAGAGATCTGATTTTCTGATTTCCGTATGTTCCGACCATTTGATCTTATGCTG 1128  
 QY 712 TATAGTATAGTCAAGACTCTCCGTGCTGCTCCAGAGGGTGTGATGCTGATAGG 771  
 Db 1129 TTGAGTGTGTTAAGACTCATCTGCGTCCGTACCAAGAGGTTTATGATTTCTGATTA 1188  
 QY 772 GTTGACCAACCAATCCGATCATTTGTTGCTGATGGAATCGAAGATGTTACATTAAGGTC 831  
 Db 1189 GTTGACACTCCCATAGAAATTAATAGTTGCTGATGGAATAGAGATGTTACATTAAGGTC 1248  
 QY 832 TCAGATGAAGTGGAGATTAACGAGAAAGGCTCCCTAAATATTAATTAATTAATTAAT 891  
 Db 1249 TCAGATGAAGGAGTGAATGCAAGAGGTTGCTTAATAATTTTACATATATATAT 1308  
 QY 892 AGCACTGCAAGAAACCACTGGAAGATGATGGAATGGAATGGAATGGAATGGAATGGAAT 951  
 Db 1309 AGTACTGCCAGAAACCACTGGAAGATGATGGAATGGAATGGAATGGAATGGAATGGAAT 1365  
 QY 952 ACTAGGCTGTTAGTATGATGCTGCTGCTTATGATGCTGCTGCTGCTGCTGCTGCTGCT 1011  
 Db 1366 ACAATGGCTGATATGATATGATGCTTCTTAATAGTCTGCTATATGCTGCTGCTGCTGCT 1425  
 QY 1012 GGAGATTGGAGATCATATCCATGATGAAGATAGGGAGTGAATGCTTACTTGCATTGCT 1071  
 Db 1426 GGTGATCTTCAAAATTAATCTTATGAGAGATATGGCACTGATGATATCTCCATTGTCT 1485  
 QY 1072 CGTCTTGAGACTCGCAGAGAGCCTTGGCATTGA 1104  
 Db 1486 CGTTGGAGATTCAACAAGAACCTTGGCTTGA 1200  
  
 RESULT 12  
 US-10-062-254-9  
 ; Sequence 9, Application US/10062254  
 ; Publication No. US20020138882A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cahoon, Edgar B  
 ; APPLICANT: Cahoon, Rebecca E  
 ; APPLICANT: Falco, Saverio Carl  
 ; APPLICANT: Fang, Yiwen  
 ; APPLICANT: Hanke, Sabine S.  
 ; APPLICANT: Lee, Jian-Ming  
 ; APPLICANT: Li, Zhongsen  
 ; APPLICANT: Miao, Guo-Hua  
 ; APPLICANT: Morgante, Michele  
 ; APPLICANT: Niu Xiping  
 ; APPLICANT: Odell, Joan  
 ; APPLICANT: Rafalski, Antoni  
 ; APPLICANT: Sakai, Hajime  
 ; APPLICANT: Zheng, Peizhong  
 ; APPLICANT: Zhu, Qun  
 ; TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism  
 ; FILE REFERENCE:  
 ; CURRENT APPLICATION NUMBER: US/10/062,254  
 ; PRIOR FILING DATE: 2002-02-01  
 ; PRIOR APPLICATION NUMBER: 09/630,346  
 ; PRIOR FILING DATE: 2000-07-28  
 ; PRIOR APPLICATION NUMBER: 60/146511  
 ; PRIOR FILING DATE: 1999-07-30  
 ; PRIOR APPLICATION NUMBER: 60/156006  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: 60/156899

/ PRIOR FILING DATE: 1999-09-30  
 / PRIOR APPLICATION NUMBER: 60/157287  
 / PRIOR FILING DATE: 1998-10-01  
 / PRIOR APPLICATION NUMBER: 60/169767  
 / PRIOR FILING DATE: 1999-12-09  
 / PRIOR APPLICATION NUMBER: 60/171054  
 / PRIOR FILING DATE: 1999-12-16  
 / PRIOR APPLICATION NUMBER: 60/172958  
 / PRIOR FILING DATE: 1999-12-21  
 / PRIOR APPLICATION NUMBER: 60/175151  
 / PRIOR FILING DATE: 1999-12-22  
 / PRIOR APPLICATION NUMBER: 60/173535  
 / PRIOR FILING DATE: 1999-12-29  
 / NUMBER OF SEQ ID NOS: 375  
 / SOFTWARE: Microsoft Office 97  
 / SEQ ID NO 9  
 / LENGTH: 1481  
 / TYPE: DNA  
 / ORGANISM: *Oryza sativa*  
 / US-10-062-254-9

Query Match 49.6%; Score 547.4; DB 13; Length 1481;  
 Best Local Similarity 69.9%; Pred. No. 6.6e-167;  
 Matches 778; Conservative 0; Mismatches 311; Indels 24; Gaps 2;

Db 4 GCGGTGAAGAGGCTTACGAGATGTTTCGAAAGCTTGAATCGAGACGTTCAACAGATGG 63  
 46 GGGGGGAGATGGCGTCGAGACCGGTGGCGAGGGCGGTGGCGAGAGAGTGGGAGATGG 105  
 Oy 64 GGATGCAATGAAGCAGACGGGGGTGAGCTTCAGTACATGATGAGTTCGGTTCACCTCC 123  
 106 GGGAGCATGAAGACGAGGGGGTGAAGCTTGGCTACATGATGAGTTCGGGTCCGCGCC 165  
 Oy 124 ACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAACAAGAGCTTCGATTCGGATC 183  
 Db 166 ACGAGGCGGAACCTGCTGCTCTCGCGCAGTTCTTGAGAGAGAGCTCCCAATCCGATC 225  
 Oy 184 GCGAGGCTGCGATCGAACTCGAGACGCTGCTTATGAGCTTCTGAGAAACCTGCGCTC 243  
 226 GCGGCGCGCGCTCGAGCTCGAGTCCCTCCCTTCGCGCTTCGCGAAGCCGCGCATC 285  
 Db 244 TTGAAGGTAAAGATGATGATGAGTGCATTCAGAGGACATGAGAGGTTCTCTGAGATC 303  
 286 CTCAGAGTGGGAGATTTGTAATCTTGAATCTTTCGCGACATCGAATCTTTCAGAGATG 345  
 Oy 304 AAGGATCTCTGATGAGAAAGTTTCAACAAGATGATCAAGGCTGTTAAAGTAAGACAC 363  
 346 AAGAAATGGAACGACGAGCTCGCTTTTACGCAATGATCAAGATGATCAAAAGTACCCAC 405  
 Db 364 AACAAAGTGTCCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 412  
 406 AATTAATGTGTTCACAGATGAGCTGAGGATTCACAGCTCAAGACGACGATATCCG 465  
 Oy 413 -GAATGAACCTTCAAGAAAGCTTGAATGAGATTCATCAAGTTTCTGATGCTCTATCTG 471  
 466 ACAAGGAAGATCCCACTGATTTGAGAGATTCACAGATTTCTTACAGATTTCTACATG 525  
 Db 472 TCTCGATAGGATCCGATGCTTATCGGCGACAGATGATGAGTTCATTAATCAACCA 531  
 526 TCAAGATTTGGATTCGCGATGCTTATAGGCGACGATGAGCTTTCATGATCTGACCCA 585  
 Oy 532 CCACCTTCAACAGTGGGTTACATACACCAAGATGCTCTTATGAGAGTTCGCAAGAT 591  
 586 GAGCTGGGCTCATAGGCTCATTAATACAGATTTCTCCCTATACAGGTGCTCAAGCT 645  
 Oy 592 GCTAGTGAAGATGCAAGTTCATTTGTTCAAGATATGTTGCTCCGAGATTAAC 651  
 646 GCGAGTGAAGATGCCGTTCTATTTGTTGAGGAGATATGATAGCTCTTGAATGAC 705  
 Oy 652 AATATAGGATTCAGATTTCTTCTGATGCTTCGATGCTTCGACCATTTGATTTATGCTG 711  
 706 AATCTATGAGACCACTTTTACATTTTCAATATGTTTCACTACATCTCATCTCATGCTA 765

Oy 712 TATGATTAAGTCAAGAACTCTCTCGGTGCTGTCGAGAGCGGTTGTTGACTCTGATAG 771  
 766 TTGGAATGCTGAAGAACTCTGCGTGGAGTACAGAGAGATATGATATCCGATTA 825  
 Db 772 GTTGCAACCCCATCCGATATGATTTGCTGATGATGAGATTTTCAATTAAGTTC 831  
 826 GATGTTCTTCAATGATTAATATGTTGCTGATGAGGAGAGATGTAATCTAAGGTT 885  
 Oy 832 TCAGATGAAGTGAAGTATACGAGAGACGGTCTCCCTAATAATTCATTAACCTTAC 891  
 886 AGTATGAAGTGTGTAATATCAAGAGAGCGGCTTCAAGATTTTCAATATCTATAT 945  
 Oy 892 AGCATCTGAAGAAACCCACTTGAAGAGATGAGACTTGGAAACGGCTGATGTTCCCTG 951  
 946 AGCATGCAAGAAATCCACT-----GATATGATTTGCCCTATGAGAGAGTA 993  
 Oy 952 ACTATGCTGTTATGATGATGCTGCTTATGAGCTGCTGATGCTGCTGATTTGCT 1011  
 994 ACTATGCTGATATGCTATGCTATGCTTTCATCAAGTGGCTGATGCTGATATTTGCT 1053  
 Db 1012 GAGATTTGCAATCATATCATGAGAGATACGAGACTGATGCTTACTTGAATGCTTCT 1071  
 1054 GGTGACCTGCAATCATCTCTATGAGAGATACGAGACGATGCTTACTGACTTGTCA 1113  
 Oy 1072 CGTCTGAGAGCTCCGAGAGGCTTTGCCATGA 1104  
 1114 CGTTGGAGATTCGAGAGGCTTTGCCCTTGA 1146  
 Db

RESULT 13  
 US-10-437-963-74020  
 / Sequence 74020, Application US/10437963  
 / Publication No. US2004012343A1  
 / GENERAL INFORMATION:  
 / APPLICANT: La Rosa, Thomas J.  
 / APPLICANT: Kovalic, David K.  
 / APPLICANT: Zhou, Yihua  
 / APPLICANT: Cao, Yongwei  
 / APPLICANT: Wu, Wei  
 / APPLICANT: Boukharov, Andrey A.  
 / APPLICANT: Barbazuk, Brad  
 / APPLICANT: Li, Ping  
 / TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
 / FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 / CURRENT APPLICATION NUMBER: US/10/437, 963  
 / CURRENT FILING DATE: 2003-05-14  
 / NUMBER OF SEQ ID NOS: 204966  
 / SEQ ID NO 74020  
 / LENGTH: 2892  
 / TYPE: DNA  
 / ORGANISM: *Oryza sativa*  
 / FEATURE:  
 / OTHER INFORMATION: Clone ID: PAT\_MRT4530\_74244C.1  
 / US-10-437-963-74020

Query Match 49.6%; Score 547.4; DB 18; Length 2892;  
 Best Local Similarity 69.9%; Pred. No. 9.8e-167;  
 Matches 778; Conservative 0; Mismatches 311; Indels 24; Gaps 2;

Oy 4 GCGGTGAAGAGGCTTACGAGATGTTTTCGAAAGCTTGAATCGAGACGTTCAACAGATGG 63  
 1139 GGGGGGAGATGGCGTCGAGACCGGTGGCGAGGGCGGTGGCGAGAGAGTGGGAGAGTGG 1198  
 Db 64 GGATGCAATGAAGCAGACGGGGGTGAGCTTCAGTACATGATGAGTTCGGTTCACCTCC 123  
 1199 GGGAGCATGAAGACGAGGGGGTGAAGCTTCAATATGATGAGTTCGGGTGCGCGCC 1258  
 Oy 124 ACTGAGAGAAACCTTCTGATCTCGGCGAGTTTCTTCAACAAGAGCTTCGATTCGATC 183  
 1259 ACGAGGCGGAACCTGCTGCTCGCGCAAGTTCTTCAAGAGAGAGCTCCCAATCCGATC 1318  
 Db 184 GCGAGGCTGCGATCGAATCTCGAGAGCGCTGCTTATGAGCTCTTCAAGAAACCTGCGCTC 243  
 Oy

```

Db      1319 GCGCGCGCGCGCTCGAGCTCGAGTCCCTCCCTCCGCGCTCCCGCAAGCCCGCATC 1378
Qy      244 TTGAAGTAAAGATTGGTATGTGAGTCAATTCAGGACATGAGAGCTTTCGATG 303
Db      1379 CTAAAGTGGGAGTGGTACTTGGACTCTTCCGCAACATCAATCTTCCAGAAAGT 1438
Qy      304 AAGGATCTGCTGATGAGAAAGTTCACACAGATGATCAAGGCTGTTAAAGTAAGGAC 363
Db      1439 AGGAATAGGAACGACGAGCTCGCTTTTACGAGATGATCAAGATCAAGTACGCCAC 1498
Qy      364 AACACGTGTTCCCATGATGCTCTGGTGTGAACCACTGAAGAAAG----- 412
Db      1499 AATTAATGTGTTCCAAAGATGCGACTGGAGTTCAACAGCTCAAGAACGAGATATGCG 1558
Qy      413 -GAATGAACTCTACGAAAAAGCTTGATGATTCATGATTCATGCTTACTG 471
Db      1559 ACAAGAAAGATCCCGCATGCAATTCAGAAATTCACAGATTTCTTGACAGATTTCAATG 1618
Qy      472 TCTCGATAGGATCCGATGCTTATCGGCGAGCATGTTGATGATTAATCCAAACCA 531
Db      1619 TCAGAGATTGGATTCGCGATGCTTATAGGCGAGATGCTTTCATGATTCGACCCA 1678
Qy      532 CCACTTCAACAAGTGGTTACATACACACCAAGATGCTCTATGAGGTGGCAAGGAT 591
Db      1679 GAGCTGGCGCTCATAGGCTCATTAATACAGAAATGTCCTATACAGGTGGCTCAAGCT 1738
Qy      592 GCTAGTGAAGATGCAAGGTGATTTGTTTCAAGAGATAGTGTGCTCCGAGATTAAC 651
Db      1739 GCCAGTGAAGATGCCCTTCTATTTGTTTGAAGGAAATAGATCAAGCTCTGAGATTTGAC 1798
Qy      652 ATATATGCGCATCCAAATTTCACTTTCCGATGTTCCGACCCATTTGATCTTATGCTG 711
Db      1799 ATCTATGAGAGACCACTTTTACATTTTCCATATGTTTCAATCAGATCTACATCTACTA 1858
Qy      712 TATGAGTATGATCAAGAACTCTCTCCGCTGCTGTCAGAGCGGTTGTGACTGTATAGG 771
Db      1859 TTTGAATGGTGAAGAACTCTTGCGCGTAGTACAGAAAGATATGAAATTCGATATAA 1918
Qy      772 GTTGCAACCAATCCGATCATTTGTTGCTGATGAGATCGAAGATGTTACAAATTAAGTCT 831
Db      1919 GATGTTCTTCCAGTTGAATTTATGTTGCTGATGAGGCAAGAGATGTAACATCAAGGTT 1978
Qy      832 TCAGATGAAGGTGAAGGTATACGAGAGAGGCTCTCCCTAAATATTCATCACTTAC 891
Db      1979 AGGATGAAAGTGTGAGATTAACAGAGAGCGGCTTCCAAAGATTTTCAATATCTATAT 2038
Qy      892 AGCAGTGAAGAAACCACTTGAAGAGATGTGAACTTGGAAACCGCTGATGTTCCCTG 951
Db      2039 AGCAGTGAAGAAATCCACT-----GATATGATGCTTGAAGAGAGTA 2086
Qy      952 ACTATAGCTGTATATGTTATGCTGCTTATAGTATGCTTATGCTATGCTATTTGCT 1011
Db      2087 ACTATAGCTGTATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATTTGCT 2146
Qy      1012 GGAGATTTGAGATCATCATCATGAGAGAGATGAGAGAGTGTGCTTACTTCTGCTTGTCT 1071
Db      2147 GGTGACCTGCAATTCATCTCTATGAGAGATACGGCACCGATGCTTACCTGCACTTGTCA 2206
Qy      1072 GGTCTTGAAGACTGCGAGAGAGCTTTGGCAATGA 1104
Db      2207 CGGTTGGAGATTCGAGAGAGAGCTTGTGTA 2239

```

RESULT 14  
US-10-425-114-15131

; Sequence 15131, Application US/10425114  
; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jindong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

```

; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OR INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OR INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 15131
; LENGTH: 1165
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3051-054-H6_FLI
US-10-425-114-15131

```

Query Match 48.6%; Score 536.6; DB 17; Length 1165;

Best Local Similarity 75.5%; Pred. No. 1.9e-163;

Matches 697; Conservative 0; Mismatches 214; Indels 12; Gaps 2;

```

Qy      191 GTGCGATCGAAGCTCGAGAGCTGCTTATGCGCTCTGTGAAACCTGCCGCTTGAGG 250
Db      2 GGGCATTGAGCTTGAGAAATCTTCCCTATGCTTGTCTCAAAAACCTGCTGTTTAAAG 61
Qy      251 TAAAGATTGATATGATGATGATCAATTCAGGACATGAGAGCTTTCGATCAAGATTA 310
Db      62 TTAGGATTTGATATGATGATTTCTTCCGATATCGAGAGCTTCCCAACATCAAGATG 121
Qy      311 CTGCTGATGAGAAAGATTCACACAGATGATCAAGCTGTTAAATGAGGACAAACG 370
Db      122 TGATGATGAGAAAGATTTCACTGAATGATGATCAAGGCAACAAAGTGAAGCACAAATG 181
Qy      371 TGGTTCCTGATGAGGCTGCTGGGTGTGAACAGCGTGAAGAAAGAAATGAAC----- 421
Db      182 TGGATCCCAAAAGGCTTGGGTGTGAACAGCGTGAAGAAAGAAATGATGATCAAGATTTG 241
Qy      422 TCTACGAAAGCTTGAATGATGATTCATGATTTCTGATGCTTCTACTGTTCTGATAG 481
Db      242 TTTATGAGATCTTGTGATGATTCATGATTTCTGATGCTTCTACTGTTCTGATAG 301
Qy      482 GATTCGATGCTTATGCGGAGATGTTGATGATTTGATTTGATTTGATTTGATTTGATTTG 541
Db      302 GAATCCGATGCTTATGCGGAGATGTTGATGATTTGATTTGATTTGATTTGATTTGATTTG 361
Qy      542 CAGTGGTATACACACCAAGATGCTCTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 601
Db      362 TTGTGGCTATATACACCAAAATGCTCTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 421
Qy      602 ATGCAAGTCTGATTTGTTTCAAGAGATGTTCTGCTCCGAGATTAACATATATGAGCG 661
Db      422 ATGCAAGTCTGATTTGTTTCAAGAGATGTTCTGCTCCGAGATTAACATATATGAGCG 481
Qy      662 ATGCAAGTCTGATTTGTTTCAAGAGATGTTCTGCTCCGAGATTAACATATATGAGCG 721
Db      482 ATGCAAGTCTGATTTGTTTCAAGAGATGTTCTGCTCCGAGATTAACATATATGAGCG 541
Qy      722 TCAAGAACTCTGCGGCTGCTCAAGAGAGGCTTGTGATCTGATGAGGTTGACACAC 781
Db      542 TTAAGAACTCTGCGGCTGCTCAAGAGAGGCTTGTGATCTGATGAGGTTGACACAC 601
Qy      782 CAATCGTATCATTTGTTGCTGATGATTCGAAGATGTTCAATTAAGTCTCAAGATGAG 841
Db      602 CAATCGTATCATTTGTTGCTGATGATTCGAAGATGTTCAATTAAGTCTCAAGATGAG 661
Qy      842 GTGAGGTATACGAGAGAGGCTCTCTTAATATTCATTAACCTTACACAGCTGCA 901
Db      662 GAGTGAAGATTCGAAGAGGCTCTCTTAATATTCATTAACCTTACACAGCTGCA 721
Qy      902 GAAACCACTTGAAGAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 961
Db      722 GAAACCACTTGAAGAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 778
Qy      962 GTTATGTTATGCTGCTTATGATGCTTGTATGCTGCTTGTATGCTGCTTGTATGCTGCT 1021

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 12, 2005, 04:37:23 ; Search time 4249 Seconds

(without alignments)  
9890.081 Million cell updates/sec

Title: US-10-642-531-1

Perfect score: 1104  
Sequence: 1 atcgcggtgagaagagctag.....cgcagagagccttgcacatga 1104

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 66479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

1: gb\_ests1.\*  
2: gb\_ests2.\*  
3: gb\_nuc.\*  
4: gb\_ests3.\*  
5: gb\_ests4.\*  
6: gb\_ests5.\*  
7: gb\_ests6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	878.4	79.6	1329	3	CNS0A6IR
2	878.4	79.6	1360	3	CNS0A4TA
3	876.8	79.4	1397	3	CNS0A6MN
4	875.2	79.3	1350	3	CNS0A6XJ
5	874.4	79.2	1386	3	CNS0A6R4
6	531.6	48.2	793	6	CD827897
7	510	46.2	1095	9	CL961615
8	501.2	45.4	1546	3	AY110322
9	488	44.2	1438	3	AY109472
10	481.6	43.6	807	5	BX839689
11	465.2	42.1	1098	6	CL973570
12	455.4	41.2	554	6	CD831928
13	450.2	40.8	784	5	BX839466
14	441.6	40.0	1090	7	CK208011
15	409	37.0	669	7	CN892416
16	408	37.0	780	7	CF209870
17	405	36.7	807	6	CB673677
18	401	36.3	730	7	CN185956
19	398.4	36.1	851	6	CB652358
20	390.8	35.4	671	1	AV783403
21	390	35.3	841	7	CP437749
22	385.8	34.9	782	7	CF201031
23	379	34.3	774	6	CB674701
24	375	34.0	833	7	CO103634

25	371.8	33.7	730	4	BI971303
26	367.4	33.3	686	7	CN876886
27	362.6	32.8	872	7	CO481703
28	362	32.8	653	7	CN915299
29	359.2	32.5	616	7	CN907157
30	358.2	32.4	1023	4	BG837999
31	356.4	32.3	628	7	CN918862
32	355.4	32.2	815	7	CN185955
33	351.4	31.8	674	6	CB008117
34	350.6	31.8	442	1	AV557206
35	350.4	31.7	621	7	CN919108
36	350.4	31.7	785	4	BJ570152
37	350	31.7	585	7	CN896098
38	342	31.0	612	2	BE346718
39	341.6	30.9	571	2	BE803392
40	341.6	30.9	585	4	BI426713
41	341.6	30.9	585	4	BI426722
42	341.6	30.9	666	7	CN907413
43	341.4	30.9	575	6	CA936210
44	338.2	30.6	592	6	CB257858
45	336.4	30.5	706	4	BI309456

## ALIGNMENTS

### RESULT 1

CNS0A6IR

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

1. 1329  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"

/clone="GSLT.S392G08"  
 /tissue\_type="Adult vegetative tissue"  
 /plasmid="pCMVSPORT\_6"  
 complement(1..1329)  
 /gene="At3g06483"

Query Match 79.6%; Score 878.4; DB 3; Length 1329;  
 Best Local Similarity 87.9%; Pred. No. 2.5e-259;  
 Matches 970; Conservative 0; Mismatches 131; Indels 3; Gaps 1;

```

Oy 1 ATGGCGGTGAAGAGGCTAGCGAGATGTTTTCGAAGAGCTTGATCGAGACCTTCAACA 60
Db 69 ATGGCGGTGAAGAGGCTAGCGAGATGTTTTCGAAGAGCTTGATCGAGACCTTCAACA 128
Oy 61 TGGGGATGATGAGAGAGCGGCGTGAAGCCCTCAGTACATGATGAGTGGTTCCTCACT 120
Db 129 TGGGGATGATGAGAGAGCGGCGTGAAGCCCTCAGTACATGATGAGTGGTTCCTCAACA 188
Oy 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGAGTTCCTTCAACAAGAGCTTCCGATTCCG 180
Db 189 CCTACTGAGAGAAACCTTCTGATCTCGGCGAGTTCCTTCAACAAGAGCTTCCGATTCCG 248
Oy 181 ATCGGAGAGCGTCCGATGAGTGAAGCTGAGAGCGTCTTATGAGCTTCTTGAAGAACTGCC 240
Db 249 GTCCGAGAGAGCGTCCGATGAGTGAAGCTGAGAGCGTCTTATGAGCTTCTTGAAGAACTGCC 308
Oy 241 GTCTTGAAGTGAAGATTTGGATGATGAGTCAATTCAGAGACATGAGAGCGTTCCTCGAG 300
Db 309 GTTTGAGAGTGGCGGAGTTGGTATTTGAGATCTTTCAGAGACATGAGAGCGTTCCTCGAG 368
Oy 301 ATCAAGGATATCTGCTGATGAGAGAGTTCACACAGATGATCAAGGCTGTAAAGTAAAG 360
Db 369 ATTAAGGATATCTGCTGATGAGAGAGTTCACACAGATGATCAAGGCTGTAAAGTAAAG 428
Oy 361 CACAACAAGTGGTTCCTGATGATGAGTTCATGATGATGATGATGATGATGATGATGATGAT 420
Db 429 CATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 487
Oy 421 CTTCACGAAAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 488 -TTCTGGAATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 545
Oy 481 GGGATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 546 GGGATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 605
Oy 541 ACAGTGGGTTACATACACACCAAGATGCTCTTATGAGAGGAGGAGAGAGATGATGAGAA 600
Db 606 ACAGTGGGTTATATACACACCAAGATGCTCTTATGAGAGGAGGAGAGATGATGAGAA 665
Oy 601 GATCGAGAGTGCATTTTTCATGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 666 GATCGAGAGTGCATTTTTCATGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 725
Oy 661 GATCGAGAGTGCATTTTTCATGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 726 GATCGAGAGTGCATTTTTCATGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 785
Oy 721 GTCAAGAGCTCTCTCGGCTGCTCAAGAGCGGTTTGTGACTCTGATAGGGTTGACCA 780
Db 786 GTCAAGAGCTCTCTCGGCTGCTCAAGAGCGGTTTGTGACTCTGATAGGGTTGACCA 845
Oy 781 CCAATCCGATCATTTTGTGCTGATGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 846 CCAATCCGATCATTTTGTGCTGATGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 905
Oy 841 GGTGAGGATATACCGAGAGCGGCTCTCCCTAAATATTCATCTTACCTTACAGACTGCA 900
Db 906 GGTGAGGATATACCGAGAGCGGCTCTCCCTAAATATTCATCTTACCTTACAGACTGCA 965
Oy 901 AGAAACCCACTTGAAGAGATGAGAGCTTGGAGACCGCTGATGATGATGATGATGATGATGAT 960

```

```

Db 966 AGAAACCCGCTTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1025
Oy 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 1026 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1085
Oy 1021 CAGATCATATTCATGAGAGAGATGAGAGCTGATGATGATGATGATGATGATGATGATGATGAT 1080
Db 1086 CAGATCATATTCATGAGAGAGATGAGAGCTGATGATGATGATGATGATGATGATGATGATGAT 1145
Oy 1081 GATCGGAGAGAGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1104
Db 1146 GATCGGAGAGAGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1169

```

RESULT 2  
 CDS047A  
 LOCUS  
 DEFINITION  
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
 GSLT.S122A07 of Adult vegetative tissue of strain col-0 of  
 Arabidopsis thaliana (thale cress).  
 BX823217  
 ACCESSION  
 Version  
 Keywords  
 SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS  
 TITLE  
 A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation  
 Unpublished  
 2 (bases 1 to 1360)

JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 Direct Submission  
 Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT  
 The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out  
 full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castelli  
 V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,  
 Schachter V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information center for Protein  
 Sequences). 5 prime and 3 prime are assembled with Phrap.  
 http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
 length  
 http://www.genoscope.cns.fr/cgi-bin/gdb/gdb?source=Arabidopsis.  
 Location/Qualifiers  
 1..1360

FEATURES  
 source  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSLT.S122A07"  
 /tissue\_type="Adult vegetative tissue"  
 /plasmid="pCMVSPORT\_6"  
 complement(1..1360)  
 /gene="At3g06483"

Query Match 79.6%; Score 878.4; DB 3; Length 1360;  
 Best Local Similarity 87.9%; Pred. No. 2.5e-259;  
 Matches 970; Conservative 0; Mismatches 131; Indels 3; Gaps 1;



QY 1 ATGGCGGTGAAGAGGCTAGAGATGTTTTGAAAGCTTGTATCGAAGACCTTCAACGA 60  
 Db 96 ATGGCAGTGAAGAAAGCTGGGAAATGTTCCGAAGGTTGATCGAAGATGTTCAAAA 155  
 QY 61 TGGGGATGCTAGAAAGCAGACGGGGCTGAGCTCAGGTCATGATGAGATTGCTTCCACT 120  
 Db 156 TGGGGTTCATGAAGCAACCGGTTGACCTTGATGATCAATGATGAGATTGCTTCCAAA 215  
 QY 121 CCCCATGAGAGAAACCTTCTGATCTGGCGCAGATTTCTTCAAGAGAGCTTCCGATTCGG 180  
 Db 216 CCACTGAGAGAAATCTTTGATTTCTGCTCAGATTTTGCATTAAGAGCTTCCGATTCGC 275  
 QY 181 ATGCGAGGCGTGGCATGAACTCGAGACGCTGCTTAATGAGCTCTGAGAAACCTGCC 240  
 Db 276 GTCCCGAGAGAGCGATCGAACTCAAGACGCTTCTTAATGATGCTCTGATTAACCTGCC 335  
 QY 241 GCTTTGAAGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
 Db 336 GTTTTGAAGTGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 395  
 QY 301 ATCAAGGATACCTGCTGATGAGAAAGATTTCAACAGATGATCAAGGCTTAAAGTAGG 360  
 Db 396 ATTAAAGATTCGGGTGACGAAAGATTTTCACTAGATGATTAAGGCTGCAAGTAGAG 455  
 QY 361 CACAACAACGTGTTCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420  
 Db 456 CATAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 514  
 QY 421 CTCTACAAAAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
 Db 515 -TTCTGAAATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 572  
 QY 481 GGGATCCGTAATGCTTATCGGGCAGCATGTTGATGATGATGATGATGATGATGATGATG 540  
 Db 573 GGGATCCGTAATGCTTATCGGGCAGCATGTTGATGATGATGATGATGATGATGATGATG 632  
 QY 541 ACAAGTGGTATACATACACCAAGATGCTCTTATGAGAGTGGCAAGAAATGCTAGTAGA 600  
 Db 633 ACAAGTGGTATACATACACCAAGATGCTCTTATGAGAGTGGCAAGAAATGCTAGTAGA 692  
 QY 601 GATGCAAGTTCGATTTTGTTCAGAGATGATGATGATGATGATGATGATGATGATGATG 660  
 Db 693 GATGCTCGGTCATTTGTTCCGAGATGATGATGATGATGATGATGATGATGATGATGATG 752  
 QY 661 GATCAAGTTCATCTTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
 Db 753 GATCCAGTTCATCTTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATG 812  
 QY 721 GTCAGAACTCTCTCCGTCGTCCTCAAGAGCGGTTTGTGATCTGATAGGTTGACCA 780  
 Db 813 GTCAGAACTCTCTAGCTGTCCTCAAGAGCGGTTTGTGATCTGATAGGTTGACCA 872  
 QY 781 CCAATCCGTCATTTGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 840  
 Db 873 CCAATCCGTCATTTGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 932  
 QY 841 GGTGAGGTATACGAGAGCGGTCCTCAAAATATTCCTTAACCTTAACAGACCTGCA 900  
 Db 933 GGTGAGGTATACGAGAGCGGTCCTCAAAATATTCCTTAACCTTAACAGACCTGCA 992  
 QY 901 AGAAACCCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960  
 Db 993 AGAAACCCGCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1052  
 QY 961 GGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
 Db 1053 GGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1112  
 QY 1021 CAGATCATTCATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
 Db 1113 CAGATCATTCATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1172  
 QY 1081 GACTCGAGAGAGCTTTGCAATGA 1104

Db 1173 GATTCGCAAGAGCTTTACCTGTA 1196  
 RESULT 3  
 CDS0A6M  
 LOCUS  
 DEFINITION  
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLT1S88ZG07 of Adult vegetative tissue of strain col-0 of Arabidopsis thaliana (thale cress).  
 ACCESSION  
 BX823886  
 VERSION  
 BX823886.1  
 KEYWORDS  
 HTC; GSRT cDNA.  
 SOURCE  
 Arabidopsis thaliana (thale cress)  
 ORGANISM  
 Arabidopsis thaliana  
 Bacteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE  
 1 (bases 1 to 1397)  
 Castellani V., Aury J.M., Jallion O., Wincker P., Clepet C., Menard M., Cruaud C., Quetier F., Scarpelli C., Schachter V., Troschier V., Weissenbach J., Salanoubat M., URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information Center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.  
 http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_Full Length  
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
 FEATURES  
 source  
 1..1397  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSLT1S88ZG07"  
 /tissue\_type="Adult vegetative tissue"  
 /plasmid="pCMVSPORT 6"  
 complement(1..1397)  
 /gene="At3g06483"  
 ORIGIN  
 Query Match 79.4%; Score 876.8; DB 3; Length 1397;  
 Best Local Similarity 87.8%; Pred. No. 7.8e-259;  
 Matches 969; Conservative 0; Mismatches 132; Indels 3; Gaps 1;  
 Db 161 ATGGCAGTGAAGAAACCTGCGAAATGTTCCGAAGGTTGATGAGATGTTCAAAA 220  
 QY 61 TGGGGATGCTAGAAAGCAGACGGGGCTGAGCTCAGGTCATGATGAGATTGCTTCCACT 120  
 Db 221 TGGGGTTCATGAAGCAACCGGTTGACCTTGATGATCAATGATGAGATTGCTTCCAAA 280  
 QY 121 CCCCATGAGAGAAACCTTCTGATCTGGCGCAGATTTCTTCAAGAGAGCTTCCGATTCGG 180  
 Db 281 CCACTGAGAGAAACCTTCTGATCTGGCGCAGATTTCTTCAAGAGAGCTTCCGATTCGG 340



QY 181 ATCGGAGGCGTGGCATCGAAGCTGCTTATGGCTCTGAGAACTGCC 240  
 DB 341 GTGCGCAGAGAGCCATCCAGACGCTCTTATGGCTCTGATTAACCTGCC 400  
 QY 241 GTCTTGAAGTGAAGATTGTATGTGAGTCACTCAGGACATGAGACGTTTCTGAG 300  
 DB 401 GTTTTGAAGTGGGATGTATTTGGAATCTTTCAGGACATGAGACATTTCTTGA 460  
 QY 301 ATCAAGATATCTGATGAGAAAGTTCAACAGATGATTAAGCTGTTAAAGTAAAG 360  
 DB 461 ATTAAGATTCGGGTGACAGAAAGATTCACTGATGATTAAGCTGTTAAAGTAAAG 520  
 QY 361 CAACAACAGTGTTCCTGATGAGTGTGCTGCTGAGTGAACAAGTGAAGAAAGTAA 420  
 DB 521 CATACAAATGTGTTCCCATGATGCTTGGGTGTTAATCAAGCTCAAGAAAGATGAA 579  
 QY 421 CTCTACGAAAAGCTGATGAGATTTATCAATTTCTTATGCTGCTTCACTTGTCTGATA 480  
 DB 580 --TTCTGGAATCTGATGAGATTCATCAATTTCTTATGCTGCTTCACTTGTCTGATA 637  
 QY 481 GGGATCCGTATGCTTATCGGCGAGATGTTGATGCTAATCCAAACCCCACTTAC 540  
 DB 638 GGGATCCGTATGCTTATCGGCGAGATGTTGATGCTAATCCAAACCCCACTTAC 697  
 QY 541 ACAATGGTATCAATCAACCAAGATGCTCTCTATGAGGTGCAAGGATGCTAGTGA 600  
 DB 698 ACAGTGGTATATACACAAAGATGCTCTCTATGAGGTGCAAGGATGCTAGTGA 757  
 QY 601 GATGCAAGTGGATTTGTTTCAAGAGATGTTCTGCTCGGAGATTAATATATATGCT 660  
 DB 758 GATGCTGCTGATTTGTTTCCGAGAGATGCTGCTGCTGCAAGGATTAATATATGCT 817  
 QY 661 GATCAAGTTCATTTTCCGATGTTCCGATGTTCCGATGTTCACTTATGAGTATGATTA 720  
 DB 818 GATCCCAATTCATTTTCCGATGTTCCGATGTTCCGATGTTCACTTATGAGTATGATTA 877  
 QY 721 GTCAAGAACTCTCTCCGCTGCTGCTCAAGAGGCTTGTGATCTGATAGGTTGACCA 780  
 DB 878 GTCAAGAACTCTCTCCGCTGCTGCTCAAGAGGCTTGTGATCTGATAGGTTGACCA 937  
 QY 781 CCAATCCGATCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
 DB 938 CCAATCCGATCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 997  
 QY 841 GGTGAGATATCCGAGAGAGGCTCTCTTAAATATTAATTAATTAATTAATTAATTA 900  
 DB 998 GGTGAGATATCCGAGAGAGGCTCTCTTAAATATTAATTAATTAATTAATTAATTA 1057  
 QY 901 AGAAACCCACTTGAAGAGATGAGAGTGGGAAACCGCTGATGTTCCCTGATATGCT 960  
 DB 1058 AGAAACCCACTTGAAGAGATGAGAGTGGGAAATGCTGATGTTCCCTGATATGCT 1117  
 QY 961 GGTATGATATGCTGCTGCTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
 DB 1118 GGTATGATATGCTGCTGCTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1177  
 QY 1021 CAGATCATATCCATGAGAGATGAGAGTGGGATGATGATGATGATGATGATGATGATGAT 1080  
 DB 1178 CAGATCATATCCATGAGAGATGAGAGTGGGATGATGATGATGATGATGATGATGATGAT 1237  
 QY 1081 GACTCGAGAGAGCTTGGCATGA 1104  
 DB 1238 GATTGCGAAGAGCTTTACCTGA 1261

RESULT 4  
 CDS0A6XJ 1350 bp mRNA linear HTC 06-FEB-2004  
 LOCUS Arabidopsis thaliana Full-length cDNA Complete Sequence from clone  
 DEFINITION GSTRPGH0204 of Hormone Treated Callus of strain col-0 of  
 Arabidopsis thaliana (thale cress).  
 ACCESSION BX825037

VERSION  
 EX825037.1 GI:42465807  
 KEYWORDS  
 HTC; GSTR CDNA  
 SOURCE  
 Arabidopsis thaliana (thale cress)  
 ORGANISM  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 REFERENCE  
 1 (bases 1 to 1350)  
 Castell, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C.,  
 Menard, M., Gruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,  
 Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.  
 Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
 A Combined Approach to Evaluate and Improve Arabidopsis Genome  
 Annotation  
 JOURNAL  
 Unpublished  
 REFERENCE  
 2 (bases 1 to 1350)  
 Genoscope.  
 Direct Submission  
 Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
 BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr  
 Web : www.genoscope.cns.fr)  
 COMMENT  
 The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out  
 full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castell  
 V., Aury J.M., Jallion O., Wincker P., Menard M., Gruaud C.,  
 Schachter V., Weissenbach J., Salanoubat M.  
 URV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information center for Protein  
 Sequences). 5 prime and 3 prime are assembled with Piprap.  
 http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
 length  
 http://www.genoscope.cns.fr/cgi-bin/gdb/gdb?source=Arabidopsis.  
 FEATURES  
 source  
 1..1350  
 Location/Qualifiers  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="COL-0"  
 /db\_xref="taxon:3702"  
 /clone="GSTRPGH0204"  
 /tissue\_type="Hormone Treated Callus"  
 /plasmid="pCWS-PORT 6"  
 complement(1..1350)  
 /gene="At3g06483"  
 ORIGIN  
 Query Match 79.3%; Score 875.2; DB 3; Length 1350;  
 Best Local Similarity 87.7%; Pred. No. 2.4e-258;  
 Matches 968; Conservative 0; Mismatches 133; Indels 3; Gaps 1;  
 QY 1 ATGGCGGTGAAGAGGCTAGGATGTTTCAAGAGCTTGAACGAGACGTTCAACA 60  
 DB 144 ATGGCAGTAAGAAAGCTGCGAAATGTTCCGAAAGTTGATCGAAGATGTTCAAAA 203  
 QY 61 TGGGATGATGAAGACAGACGGGCGTGAAGCTCAGTACATGATGAGAGTTCGTTCCACT 120  
 DB 204 TGGGATGATGAAGACAGACGGGCGTGAAGCTTATGATATGATGAGAGTTCGTTCCAAA 263  
 QY 121 CCGACTGAGAAACCTTTCATCTCGCGCGAGTTTCTTCAACAAGAGCTTCGATTGCG 180  
 DB 264 CTAATGAGAGAACTTTCATCTTCGCTCAAGTTTTCATTAAGAGCTTCGATTGCG 323  
 QY 181 ATCGGAGGCGTGCATGCACTGAGACGCTGCTTATGAGCTCTTGAAGAAACCTGCC 240  
 DB 324 GTCCGCAAGAGAGGATCGAATCTCAGACCTCTTATGATGATGATGATGATGATGATGAT 383  
 QY 241 GTCTTGAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 DB 384 GTTTTGAAGGTCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 443  
 QY 301 ATCAAGATATCTGCTGATGAGAAAGTTCAACAGATGATCAAGGCTTTAAGTAAG 360

Db 444 ATTAAGGATTCGGGTGACGAGAAAGATTTCATCAGATGATTAAGGCTGTCAAAAGTAAAG 503  
 361 CAAACAACTGGTTCCTCCATGATGCTCTGGGTGTAACAGCTGGAAGAAAGAAATAA 420  
 Db 504 CATAACAAATGTGGTCCCATGATGAGCTTGTGTATATACGTCGAAGAAAGAAATAA- 562  
 421 CTCTACGAAAAGCTGTAGATTCATCAGTTCTTGTATGCTCTTCTACTGCTGATA 480  
 Db 563 --TTCTGAAATCTTATGATGATTCATCAGTTCTTGTATGCTCTTCTACTGCTGATA 620  
 481 GGGATCCGATGCTTATCGGACAGATGTTGATGATTAATCCAAACCAACCTTAC 540  
 Db 621 GGGATCCGATGCTTATCGGACAGATGTTGATGATTAATCCAAACCAACCTTAC 680  
 541 ACAGTGGGTACATACACCAAGATGCTCTTATGATGATGATGATGATGATGATGAT 600  
 Db 681 ACAGTGGGTACATACACCAAGATGCTCTTATGATGATGATGATGATGATGATGAT 740  
 601 GATGCAAGTGCATTTGTTCAAGAGTATGTTCTCTCCGAGATTAACATATATGAC 660  
 Db 741 GATGCTCGTCATTTGTTCTCCAGAGTACGTTCTCCACCCGAAATTAACATATATGAC 800  
 661 GATCCAAATTCATTTCCGATGTTCCGATGTTCCGATGTTCCGATGTTCCGATGTT 720  
 Db 801 GATCCCAATTCATTTCCGATGTTCCGATGTTCCGATGTTCCGATGTTCCGATGTT 860  
 721 GTTAAGAACTCTCTCCGATGTTCCGATGTTCCGATGTTCCGATGTTCCGATGTT 780  
 Db 861 GTTAAGAACTCTCTCCGATGTTCCGATGTTCCGATGTTCCGATGTTCCGATGTT 920  
 781 CCAATCCGATTCATTTGTTCAAGAGTATGTTCTCTCCGAGATTAACATATATGAC 840  
 Db 921 CCAATCCGATTCATTTGTTCAAGAGTATGTTCTCTCCGAGATTAACATATATGAC 980  
 841 GGTGAGGTATACCGAGAGCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900  
 Db 981 GGTGAGGTATACCGAGAGCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1040  
 901 AGAAGCCCACTTAAAGAGTGTGAGCTTGGGAACCGCTATGTTCCCGATGATGCT 960  
 Db 1041 AGAAGCCCACTTAAAGAGTGTGAGCTTGGGAACCGCTATGTTCCCGATGATGCT 1100  
 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
 Db 1101 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1160  
 1021 CAGATCATATCCATGAGAGATACGGGATGATGCTTACTTGCATTTGCTCTCTTGA 1080  
 Db 1161 CAGATCATATCCATGAGAGATACGGGATGATGCTTACTTGCATTTGCTCTCTTGA 1220  
 1081 GACTCCGAGAGCGCTTGGCATGA 1104  
 Db 1221 GATTCCGAGAGCGCTTACCTGA 1244

RESULT 5  
 LOCUS CDS046R4 1386 bp mRNA linear HTC 06-FEB-2004  
 DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
 GSLTPGH56ZC08 of Hormone Treated Callus of strain col-0 of  
 Arabidopsis thaliana (thale cress).  
 ACCESSION BX824559.1 GI:42465670  
 VERSION 1  
 KEYWORDS HTC; GSUT; cDNA.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Bacteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 1386)  
 Castelli, V., Aury, J.-M., Jallion, O., Winkler, P., Clepet, C.,  
 Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachner, V.,  
 Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.

TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
 A Combined Approach to Evaluate and Improve Arabidopsis Genome  
 Annotation  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1386)  
 AUTHORS Genoscope.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage ;  
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 COMMENT The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out  
 full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castelli  
 V., Aury J.-M., Jallion O., Winkler P., Menard M., Cruaud C.,  
 Schachner V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information center for Protein  
 Sequences). 5 prime and 3 prime are assembled with Phrap.  
 http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
 length  
 http://www.genoscope.cns.fr/cgi-bin/g9b/g9b?source=Arabidopsis.  
 FEATURES  
 source 1..1386  
 /organism="Arabidopsis thaliana"  
 /mol\_type="cDNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSLTPGH56ZC08"  
 /issue\_type="Hormone Treated Callus"  
 /plasmid="pCMVSPORT 6"  
 complement(1..1386)  
 /gene="At3g06483"  
 ORIGIN  
 Query Match 79.2%; Score 874.4; DB 3; Length 1386;  
 Best Local Similarity 87.8%; Pred. No. 4,3e-258;  
 Matches 966; Conservative 0; Mismatches 131; Indels 3; Gaps 1;  
 1 ATGCGGTGAAGAGGCTACGAGATGTTTTCGAAGAGCTTTCGAGAGCTTTCACAGA 60  
 Db 134 ATGCGGTGAAGAGGCTACGAGATGTTTTCGAAGAGCTTTCGAGAGCTTTCACAGA 193  
 61 TGGGATGATGAGAGAGAGCGGCGTGAAGTACGATGATGATGATGATGATGATGATG 120  
 Db 194 TGGGATGATGAGAGAGAGCGGCGTGAAGTACGATGATGATGATGATGATGATGATG 253  
 121 CCCACTGAGAGAAACCTTCGATCTCGGCGAGTTCTTCACAGAGAGCTTCCGATTCGG 180  
 Db 254 CTTACTGAGAGAAATCTTTGATTTCTGCTCACTTTTGCATAGAGAGCTTCCGATTCGG 313  
 181 ATGCGAGGCGTGCATGAGTACGAGAGCTTCCTTATGCTCTCTGAGAAACCTGCC 240  
 Db 314 GTCGCGAGAGAGCGATGCAACTCCAGAGCTTCTTATGCTCTCTGATTAACCTGCC 373  
 241 GTCCTTAAGGTAAGATGTTGTTGAGTCAATTCAGGAGCAATGAGAGCTTCCGAG 300  
 Db 374 GTTTTAAGGTTGCGGATGTTGTTGAGTCAATTCAGGAGCAATGAGAGCTTCCGAG 433  
 301 ATCAAGATATGCTGATGAGAGAGTTTCAACAGATGATCAAGAGCTTTAAAGTAAG 360  
 Db 434 ATTAAGATATGCTGATGAGAGAGTTTCAACAGATGATCAAGAGCTTTAAAGTAAG 493  
 361 CACACAACTGTTTCCATGATGCTCTGGGTGTAACAGCTGAAGAAAGAAATGAAA 420  
 Db 494 CATAACAAATGTGTTCCCATGATGCTTGGGTGTAACAGCTGAAGAAAGAAATGAAA- 552  
 421 CTCTACGAAAAGCTGTAGATTCATCAGTTCTTGTATGCTCTTACTGCTGATA 480  
 Db 553 --TTCTGAAATCTTATGATGATTCATCAGTTCTTGTATGCTCTTACTGCTGATA 610  
 481 GGGATCCGATGCTTATCGGACAGATGTTGATGATTAATCCAAACCAACCACTTAC 540

Db 611 GGGATCCGATGCTTATGCGGAGCAGCTTGATGTCATATCCAAATCACCCTTCAT 670  
Qy 541 ACGTGGGTATACATACACCAAGATGTCCTATGAGGTGCGAAGATCTAGTAA 600  
Db 671 ACGTGGGTATATACACCAAGATGTCCTATGAGGTGCGAAGATCTAGTAA 730  
Qy 601 GATCGAAGTGTGATGTTGTTACAGAGTATGTTGTCGTCGCGAATTAACATATATG 660  
Db 731 GATGCTCGGTCAATTTGTTCCGAGAGTACGTTCTGACCGGAATTAACATATATG 790  
Qy 661 GATCCAAAGTTTCACTTTCCGATGTTCCGACCCATTTGATCTTATGATGATGATTA 720  
Db 791 GATCCCAAGTTTCACTTTCCGATGTTCCAGAGCATTTGATCTTATGATGATGATTA 850  
Qy 721 GTCAAGAACTCTCCGATGTTCCAGAGCGGTTGTTGATCTGATAGGTTGACCA 780  
Db 851 GTCAAGAACTCTCTAGTGTCTGTCAGAGCGATTTGTTGATCTGATAGGTTGACCA 910  
Qy 781 CCAATCCGATCATTTGCTGATGATGAAATCGAAGATGTTAACAATAAGTCTCAGATGA 840  
Db 911 CCAATCCGATCATTTGCTGATGATGAAATCGAAGATGTTAACAATAAGTCTCAGATGA 970  
Qy 841 GGTGAGGATATCCGAGAGCGGCTCCCTAAATATTCATCTCTACAGACATGCA 900  
Db 971 GGTGAGGATATGCAAGAGCGGCTCTCCAGAAATATTCACCTTATCAGACATGCA 1030  
Qy 901 AGAAACCACTTGAAGAGATGTCGATGGAACCGCTGATGTTCCCTGATCATGATGCT 960  
Db 1031 AGAAACCGCTTGAAGAGATGTCGATGGAACCGCTGATGTTCCCTGATCATGATGCT 1090  
Qy 961 GGTATGATGATGCTGCTGCTATGATGCTGATGCTGATGCTGATGCTGATGCTGATG 1020  
Db 1091 GGAATATGATGATGCTGCTGCTATGATGCTGATGCTGATGCTGATGCTGATGCTGATG 1150  
Qy 1021 CAGATCATATCCATGAGAGATGATGAGGATGATGCTGATGCTGATGCTGATGCTGATG 1080  
Db 1151 CAGATCATATCCATGAGAGATGATGAGGATGATGCTGATGCTGATGCTGATGCTGATG 1210  
Qy 1081 GACTCGCAGAGACCTTTGCC 1100  
Db 1211 GATTGCGAAGAGCCTTACC 1230

RESULT 6  
CD827897 793 bp mRNA linear EST 10-JUL-2003  
LOCUS BN25.069A22F020124 BN25 Brassica napus cDNA clone BN25069A22, mRNA  
DEFINITION  
CD827897  
CD827897  
CD827897.1 GI:32509837  
KEYWORDS  
SOURCE  
ORGANISM  
Brassica napus (rape)  
Brassica napus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 793)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Genoplane, a major partnership french program in plant genomics  
Contact: Genoplane  
Genoplane  
93, rue Henri Rochefort 91025 EVRY CEDEX France  
Tel: 33 1 69 47 54 00  
Fax: 33 1 69 47 54 10  
This sequence has been generated in the framework of the french  
plant genomics programme 'Genoplane' (<http://www.genoplane.com>  
and <http://genoplane-info.infobiogen.fr>).  
Location/Qualifiers  
1..793  
/organism="Brassica napus"  
/mol\_type="mRNA"

ORIGIN  
/cultivar="Jet Neuf"  
/db\_xref="taxon:3708"  
/clone="BN25069A22"  
/issue\_type="seed"  
/clone\_id="BN25"

Query Match 48.2%; Score 531.6; DB 6; Length 793;  
Best Local Similarity 88.6%; Pred. No. 2.8e-152;  
Matches 576; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 24 GATGTTTGAAGAGGTTGATGAGACCTTCAAGATGAGATGATGATGATGATGATGATGATG 83  
Db 144 GATATTTCTAGAGAGGTTGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 203  
Qy 84 CCGAGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 143  
Db 204 CGTTAGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 263  
Qy 144 CTGCGCGAGTTTCTTCAAGAGCTTCCGATTCGATTCGATTCGATTCGATTCGATTCGAT 203  
Db 264 CTCGCGCAGTTTCTTCAAGAGCTTCCGATTCGATTCGATTCGATTCGATTCGATTCGAT 323  
Qy 204 CGAGACGCTGCTTATGAGCTCTCTGAGAAACCTGCTGATGATGATGATGATGATGATGAT 263  
Db 324 CGAGACGCTGCTTATGAGCTCTCTGAGAAACCTGCTGATGATGATGATGATGATGATGAT 383  
Qy 264 TGTGAGTCAATCAGAGCATGAGAGCTTCTGATGATGATGATGATGATGATGATGATGAT 323  
Db 384 TGTGAGTCAATCAGAGCATGAGAGCTTCTGATGATGATGATGATGATGATGATGATGAT 443  
Qy 324 AGAGTTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 383  
Db 444 GAGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 503  
Qy 384 GAGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 443  
Db 504 GAGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 563  
Qy 444 TCATATGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 503  
Db 564 TCATATGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 623  
Qy 504 GCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 563  
Db 624 GCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 683  
Qy 564 GATGCTCTCTATGAGAGTGGAGAGATGCTAGTGAATGCAAGTCTGATGATGATGATGAT 623  
Db 684 GATGCTCTCTATGAGAGTGGAGAGATGCTAGTGAATGCAAGTCTGATGATGATGATGAT 743  
Qy 624 AGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 673  
Db 744 AGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 793

RESULT 7  
CU961615 1095 bp DNA linear GSS 21-SEP-2004  
LOCUS OsIFCC006652 Oryza sativa Expressed Library Oryza sativa (indica  
DEFINITION  
CU961615  
CU961615  
CU961615.1 GI:52377973  
KEYWORDS  
SOURCE  
ORGANISM  
Oryza sativa (indica cultivar-group)  
Oryza sativa (indica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 1095)  
REFERENCE  
AUTHORS  
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,  
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,  
Mong, G. K. S., Deng, X. W., and Wang, J.

TITLE An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis  
JOURNAL Unpublished (2004)  
COMMENT Contact: Chen Chen  
Department of Bioinformatics  
Beijing Institute of Genomics  
Chinese Academy of Sciences, Beijing 101300, China  
Tel: 86-10-80481559  
Fax: 86-10-80488676  
Email: chenchen@genomics.org.cn  
Rice genomic sequence.  
Class: exon-trapped.

## FEATURES

source 1. .1095  
Location/Qualifiers  
/organism="Oryza sativa (indica cultivar-group)"  
/db\_xref="taxon:39946"  
/clone\_id="Oryza sativa Expressed Library"  
/note="Oryza sativa exon trapped genomic sequences"

## ORIGIN

Query Match 46.2%; Score 510; DB 9; Length 1095;  
Best Local Similarity 68.4%; Pred. No. 1.5e-145;  
Matches 746; Conservative 0; Mismatches 320; Indels 24; Gaps 2;

Oy 22 GAGATGTTTCGAAGAGCTTGATGAGAGAGCTTCAAGATGGGATGATGAAGCAGACG 81  
Db 13 GACCCGCGCGGAGGCGGCGGAGAGAGTGGCGGAGGAGATGAGCAGACG 72  
Oy 82 GCGGTGAGCTTCAGATGATGATGAGTGGTTTCACTCCACCTGAGAGAACTTTCTG 141  
Db 73 GCGGTGAGCTTCAGATGATGATGAGTGGTTTCACTCCACCTGAGAGAACTTTCTG 132  
Oy 142 ATCTGCGGAGTTCCTTCAAGAGAGCTTCCATTCGGATTCGGAGAGCTTGGATGAG 201  
Db 133 CCGTCCGCGAGTTCCTTCAAGAGAGCTTCCATTCGGATTCGGAGAGCTTGGATGAG 192  
Oy 202 CTGAGAGAGCTTCTTATGAGCTTCTCTGAGAGAACTTCCGTTTGAAGATGAG 261  
Db 193 CTGAGAGAGCTTCTTATGAGCTTCTCTGAGAGAACTTCCGTTTGAAGATGAG 252  
Oy 262 TATGAGAGTTCAGAGAGATGAGAGCTTCTTCTGAGATCAAGATGATGAG 321  
Db 253 TATGAGAGTTCAGAGAGATGAGAGCTTCTTCTGAGATCAAGATGATGAG 312  
Oy 322 AAGAGTTCAGAGAGATGAGAGCTTCTTCTGAGATCAAGATGATGAG 381  
Db 313 CTGAGATTCAGAGAGATGAGAGCTTCTTCTGAGATCAAGATGATGAG 372  
Oy 382 ATGAGCTTGGGATGAGAGCTTCTTCTGAGATCAAGATGATGAG 429  
Db 373 ATGAGCTTGGGATGAGAGCTTCTTCTGAGATCAAGATGATGAG 412  
Oy 430 AAGCTTGAAGATTCATCAGATTTCTTCTGATGCTTCTCTGATGAGATGAG 489  
Db 433 GCGATGATGAGATTCATCAGATTTCTTCTGATGCTTCTCTGATGAGATGAG 492  
Oy 490 ATGCTTATCGGAGAGATTTGAGTTCATTCACCAACCACCTTCAACAGTGGT 549  
Db 493 ATGCTTATCGGAGAGATTTGAGTTCATTCACCAACCACCTTCAACAGTGGT 552  
Oy 550 TACATACACAGAGATGCTCTCTATGAGAGTGGAGAGAGATGAGATGAG 609  
Db 553 CTCTATAGCAAGAGATGCTCTCTATGAGAGTGGAGAGAGATGAGATGAG 612  
Oy 610 TCGATTTGTTTCAGAGAGATGCTCTCTATGAGAGTGGAGAGAGATGAGATGAG 669  
Db 613 GCTATTTGATGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 672  
Oy 670 TTGATTTTCCGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 729  
Db 673 TTGATTTTCCGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 732

Oy 730 TCTCTCCGCTGCTCCAGAGAGCGTTTGTGACTGATGAGGTTGACCAACATCCGT 789  
Db 733 TCCCTCCGCTGCTCCAGAGAGCGTTTGTGACTGATGAGGTTGACCAACATCCGT 792  
Oy 790 ATCATTTGCTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 849  
Db 793 ATCATTTGCTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 852  
Oy 850 ATACCGAGAGAGCGTCTCCCTAAATATCTTACTTACTTACTTACTTACTTACTTACTT 909  
Db 853 ATACCGAGAGAGCGTCTCCCTAAATATCTTACTTACTTACTTACTTACTTACTTACTT 912  
Oy 910 CTGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 969  
Db 913 CTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 960  
Oy 970 TATGATGCTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1029  
Db 961 TATGAGATGCTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1020  
Oy 1030 TCCATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1089  
Db 1021 TCTATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1080  
Oy 1090 GAGCCTTTGC 1099  
Db 1081 GAGCCTTTGC 1090

RESULT 8  
AY110322 1546 bp mRNA linear HTC 17-OCT-2002  
LOCUS Zea mays CL1703\_1 mRNA sequence.  
DEFINITION AY110322  
ACCESSION AY110322  
VERSION AY110322.1 GI:21214656  
KEYWORDS HTC.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
Clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 1546)  
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., White, M.S.,  
Arthur, L.M., Hanafey, M., Morgante, M. and Tingey, S.V.  
Maize Mapping Project/DuPont Consensus Sequences for Design of  
Overgo Probes  
Unpublished (2002)  
2 (bases 1 to 1546)  
Coe, E.H.  
Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA  
Direct Submission  
Coe, E.H.  
Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
If you are interested in getting corresponding physical clones,  
these are publicly available from ZmDB and may be found by BLAST  
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,  
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
Schubert, Iowa State, then clones may be requested from ZmDB:  
www.zmdb.iastate.edu.

## FEATURES

source

1. .1546  
Location/Qualifiers  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="taxon:39946"  
/db\_xref="MaizeDB:630518"  
/clone\_id="Maize Mapping Project/DuPont Consensus  
Library"  
/note="this sequence is part of a project of EST  
assemblies resulting from the application of public  
contigs to seed DuPont contigs; this resource was  
assembled by DuPont as part of a collaboration for the  
overgo addressing of BACs in conjunction with the Maize  
Mapping Project"

## ORIGIN

Query Match 45.4%; Score 501.2; DB 3; Length 1546;  
 Best Local Similarity 65.0%; Pred. No. 8.7e-143;  
 Matches 718; Conservative 0; Mismatches 366; Indels 21; Gaps 2;

```

QY 12 GAAGGCTACGAGATGTTTTCAGAGAGCTTGATCGAGAGCTTCACAGATGGGAGTATG 71
DB 77 GATGGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 136
QY 72 GAAGAGAGGGGCGTGAAGCTCAGGTACATGATGAGATTGCGTTCACCTCCACTGAGAG 131
DB 137 GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 196
QY 132 AAACCTTCTGATCTCGGCGAGATTCTTTCACAGAGAGCTCGATTCGATTCGAGAGCG 191
DB 197 NNNNNNGCTCTCTCCGCGAGATTCTTGCAAGAGAGTCCCATCCGCTTCGCGCGCG 256
QY 192 TGGCATCGAAGCTGAGAGCGTGCCTTATGCGCTCTGAGAGAACTGCGCGCTTGAAGT 251
DB 257 CGCGCTCGAGCTCGAGCTGCGCTGCGCTTCGCGCTTCGAGAGAGAGAGAGAGAGAG 316
QY 252 AAGAGATTGATGATGAGAGATCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 311
DB 317 GCGGAGCTGATCTTGAAGCTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 376
QY 312 TCGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 371
DB 377 GAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 436
QY 372 GGTCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 431
DB 437 GGTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 496
QY 432 GCTT-----GATGAGATTGATGATGATGATGATGATGATGATGATGATGATGAT 479
DB 497 GGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 556
QY 480 AGGAGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 539
DB 557 TGGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 616
QY 540 CACAGTGGGTTACATACACACAGAGATGCTCCATGAGAGAGAGAGAGAGAGAGAGAT 599
DB 617 TGTATAGAGCTCATCATACATACATACATACATACATACATACATACATACATACAT 676
QY 600 AGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
DB 677 AGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 736
QY 660 CGATCCAAAGTTTCACTTTCCGATGATGATGATGATGATGATGATGATGATGATGAT 719
DB 737 CGATCCAAATTTTCACTTTCCGATGATGATGATGATGATGATGATGATGATGATGAT 796
QY 720 AGTCAGAGATCTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
DB 797 GGTGAGAGATCTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 856
QY 780 ACCAATCCGATCTCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 839
DB 857 NNNGGTTAAGATTAATGCTGATGAGAGAGAGATGATGATGATGATGATGATGATGAT 916
QY 840 AGGTGAGATTAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 899
DB 917 AGGTGAGAGATTAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 976
QY 900 AAGAAACCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 959
DB 977 AAGAAATCTCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1027
QY 960 TGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1019
DB 1028 TGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087

```

QY 1020 GCAGATCATATCCATGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1079  
 DB 1088 GCAATATCATCTCATGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1147  
 QY 1080 AGACTCGAG 1104  
 DB 1148 AGATTCAG 1172

RESULT 9  
 AY109472 1438 bp mRNA linear HTC 17-OCT-2002  
 LOCUS  
 DEFINITION Zea mays CIL1702\_1 mRNA sequence.  
 ACCESSION AY109472  
 VERSION AY109472.1 GI:21213205  
 KEYWORDS  
 SOURCE HTC.  
 ORGANISM Zea mays  
 Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 1438)  
 Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whiteside, M.S.,  
 Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.  
 Maize Mapping Project/Dupont Consensus Sequences for Design of  
 Overgo Probes  
 Unpublished (2002)  
 2 (bases 1 to 1438)  
 Coe, E.H.

REFERENCE  
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of  
 Misouri, Columbia, MO 65211, USA  
 AUTHORS If you are interested in getting corresponding physical clones,  
 these are publicly available from ZmDB and may be found by BLAST  
 searching at MSU, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,  
 www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
 Schnable, Iowa State, then clones may be requested from ZmDB:  
 www.zmdb.iastate.edu.

## FEATURES

source  
 1..1438  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4577"  
 /db\_xref="taxon:4577"  
 /clone\_lib="Maize Mapping Project/Dupont Consensus  
 Library"  
 /note="this sequence is part of a project of EST  
 assemblies resulting from the application of public  
 contigs to seed Dupont contigs; this resource was  
 assembled by Dupont as part of a collaboration for the  
 overgo addressing of BACs in conjunction with the Maize  
 Mapping Project"

## ORIGIN

Query Match 44.2%; Score 488; DB 3; Length 1438;  
 Best Local Similarity 63.5%; Pred. No. 1e-138;  
 Matches 700; Conservative 0; Mismatches 378; Indels 24; Gaps 2;

```

QY 15 GGTACGAGATGTTTTCGAGAGAGCTTGATCGAGAGAGCTTCACAGATGGGATGATGAA 74
DB 57 GCGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 116
QY 75 GCAGACGGGCGTGAGCTCAGGTACATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 134
DB 117 CCAAGACGGGCTGAGCTGCGCTACATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 176
QY 135 CTTTCTGATCTCGGCGAGATTTCTTCAAGAGAGCTTCGATTCGATCGAGAGAGCTGC 194
DB 177 CTTCTCTCGCGCGAGAGATTTCTTCAAGAGAGAGCTTCGATTCGATCGAGAGAGCTGC 236
QY 195 GATGAACTGAGAGAGCTGCTTATGAGCTCTCTGAGAGAACTGCGCTTGAAGGTTAAG 254

```

[illegible]

	ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsids.
	REFERENCE	1 (bases 1 to 807)
	AUTHORS	Castelli V., Aury J.M., Jallion O., Wincker P., Clepet C., Menard M., Cruaud C., Quetier F., Scarpelli C., Schachter V., Temple G., Caboche M., Weissenbach J. and Salanoubat M. Whole Genome Sequence Comparisons and "Full-Length" cDNA Sequences A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation Unpublished (2004)
	JOURNAL	Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). <a href="http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST">http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST</a> <a href="http://www.genoscope.cns.fr/cgi-bin/gsb/gsb?source=Arabidopsis">http://www.genoscope.cns.fr/cgi-bin/gsb/gsb?source=Arabidopsis</a> .
	FEATURES	Location/Qualifiers 1..807 /organism="Arabidopsis thaliana" /mol_type="mRNA" /ecotype="Col-0" /db_xref="taxon:3702" /closure="GSJLTS93ZE08" /tissue_type="Adult vegetative tissue" /clone_Type="Arabidopsis thaliana Adult vegetative tissue Col-0"
	ORIGIN	
Query Match	43.6%; Score 481.6; DB 5; Length 807;	
Best Local Similarity	84.0%; Pred. No. 7.8e-137;	
Matches 568; Conservative	0; Mismatches 104; Indels 4; Gaps 2;	
Ds	1 ATGCGGTGAAGAAGGGCTAGCGAGATGTTTCGAAGACCTTGATCGAGACGTTCAACA 60	
Ds	132 ATGCAGTGAAGAAGAGCCTGCGAATTTCCGAAGATTGATCGAAGATGTTCAACA 191	
Ds	61 TGGGATGATGAAAGCACAGCGGCGTAGACTTCAGTAATCATGATGAGTTCGTTCCACT 120	
Ds	192 TGGGGTTGCATGAAGCAAACCGGCTTAAAGCTTATGATAATCATGATGAGTTCGTTCCA 251	
Ds	121 CCCACTGAGAAACCTTCGATCTCGGCCGAGTTTCTTCAAGGAGCTTCGATTCGG 180	
Ds	252 CTATCTAGAGGAATCTT-TGATTTCTGCTAGTTTTGCAATTAAGAGCTTCGATTCGC 310	
Ds	181 ATCCGAGGCGGTGCGATCGAACTGAGACGCTGCTTATGGCTCTCTGAGAAACCTGCC 240	
Ds	311 GTCCCGAGAGAGGATCGAGCTCCACACGCTTCCTTATGCTCTCTGATTAACCTGCC 370	
Ds	241 GTCTTGAAGGTAGAGATTGGATGTGAGAGTCATTCAGGACATGAGAGCTTCTCGAG 300	
Ds	371 GTTTTGAAGGTGCGGAGTTGGTATTATTTGGAATCTTTCAGGACATGAGAGCTTCGAG 430	
Ds	301 ATCAAGATCTCGCTGTGAGGAAGAGTCAACAGATGATCAAGGCGCTTAAAGTAGG 360	
Ds	431 ATTAAAGATTCGGGTGACGAAAAGATTTCACTCAGATGATTAAGGCTGTAAAGTAGG 490	
Ds	361 CACAACACGTGTGTCCCATGATGCTCTGGGTGTAACACAGCTGAAGAAAGATGAA 420	
Ds	491 CATACATATGTGTTCCTCATGATGCTTTGGGTATTATACAGCTCAAGATAGGAATGA-- 548	
Ds	421 CTTAGAAAAAGCTTGAATGAGATTCATCAAGTTCTTATGCTCTTACTGTCTCGATA 480	



Db 549 - GTTCTGAAATCTTATAGATTCATCATGTTTCTTATGATCTTTCTACTGTCGGAATC 607  
 Qy 481 GGGATCCGATATGTTATCGGCGACGATGTTGATGCATATCCAAACCCACCACTTAC 540  
 Db 608 GGGATCCGATATGTTATCGGCGACGATGTTGATGCATATCCAAACCCACCACTTAC 667  
 Qy 541 ACAGTGGGTTATACACACCAAGATGTCCTATGAGGTCGCAAGAAATGCTGTA 600  
 Db 668 ACAGTGGGTTATATTTCTCAAAAGATGTCCTATGAGGTCGCAAGAAATGCTGTA 727  
 Qy 601 GATCGAAGTGCATTTGTTTTCAGAGATATGCTGCTCCGAGATTAACATATATG 660  
 Db 728 GATGCTCGGTCATTTGTTTTCAGAGATATGCTGCTCCGAGATTAACATATATG 787  
 Qy 661 GATCCAGTTTCACTT 676  
 Db 788 GATCCAGTTTCCCCCTT 803

RESULT 11  
 CL973570 1098 bp DNA linear GSS 21-SEP-2004  
 LOCUS CL973570  
 DEFINITION OsIFCC024520 Oryza sativa Expressed Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.  
 ACCESSION CL973570  
 VERSION CL973570.1 GI:52401665  
 KEYWORDS GSS.  
 SOURCE Oryza sativa (indica cultivar-group)  
 ORGANISM Oryza sativa (indica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE  
 AUTHORS Ma, L., Manger, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G. K. S., Deng, X. W. and Wang, J.  
 TITLE An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: Chen Chen  
 Department of Bioinformatic  
 Beijing Institute of Genomics  
 Chinese Academy of Sciences, Beijing 101300, China  
 Tel: 86-10-80481559  
 Fax: 86-10-80488676  
 Email: chenchen@genomics.org.cn  
 Rice genomic sequence.  
 Class: exon-trapped.

## FEATURES

## source

1..1098  
 location/Qualifiers  
 /organism="Oryza sativa (indica cultivar-group)"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:39946"  
 /clone\_lib="Oryza sativa Expressed Library"  
 /note="Oryza sativa exon trapped genomic sequences"

## ORIGIN

Query Match 42.1%; Score 465.2; DB 9; Length 1098;  
 Best Local Similarity 67.4%; Pred. No. 1e-131;  
 Matches 728; Conservative 0; Mismatches 298; Indels 54; Gaps 3;

Qy 13 AAGCTGAGGAGATGTTTGAAGAGCTTATGAGAGCTTACAGATGGGAGATCATG 72  
 Db 1 ATGGCTGCGAGCGCGGCGAGCGCGTGGCGAGAGAGTGGGAGTGGGAGCATG 60  
 Qy 73 AAGCAGACGGGCGCTGAGCTCAGGTACATGATGAGTTCGCTCCACTCCACTGAGAA 132  
 Db 61 AAGCAGACGGGCGGAGGCTCGCTCATCATATGAGATTCGGTCCGCGCCCAAGGAGCG 120  
 Qy 133 AACCTTCTGATCTTCGGGCGCATTTCTTCAACAAGACCTTCCGATTTGGATTCGAGGCGT 192  
 Db 121 AACCTCTGCTCTCCGCGGAGTTCCTGAGAGAGAGACTCCCAATCCGAGATCCGCGCGCC 180

Qy 193 GCGATCGAACTCGACAGCGCTGCTTATGCGCTCTCTGAGAAACCTGCGCTTGAAGATA 252  
 Db 181 GCGCTGAGCTCGAGTCCCTCCCTTCCGCTCTTCCCGAAGCCGCACTCTCAAGATA 240  
 Qy 253 A-----GAGATTGTATGAGTGCATTCAGGGAC 282  
 Db 241 ATTCTGGGGAGTGAAGTGTGAATAGTCCGAGATTGGTACTTGGACTCTTCCGCGAC 300  
 Qy 283 ATGAGACGCTTCTCTGAGATCAAGATATGCTGATGAGAAAGATTCACAGATGATC 342  
 Db 301 ATGAGATCTTCTCAGAAAGTGAAGATTAAGAAAGACAGACTGCTTTTATGCAAGATATC 360  
 Qy 343 AAGCGTTTAAAGTAAAGCAACAACGCTGCTCCATGATGATGCTTGGGTGAAACAG 402  
 Db 361 AAGATGATCAAGTATAGCCCAATATATGTTTCCAAACATGGCACTGGGATTCACAG 420  
 Qy 403 CTGAAGAAG-----GAATGAACCTTACGAAAGCTGATGAGATTATCAG 450  
 Db 421 CTCAGAGACGACGATATGCAACAAGAAAGATCCCACTGCAATTGACGAGATCCACAG 480  
 Qy 451 TTTCTTGATCGCTTCTACTTGTCTGTATAGGATCCGATGCTTATCGGCGAGCATGT 510  
 Db 481 TTTCTTGACAGATTTCTACATGTCAGAGATGGGATCCGATGCTTATAGGCGAGCATGT 540  
 Qy 511 GAGTTCATATATCCAAACCCACTTCAACAGTGGGTTACATACACCAAGATGCT 570  
 Db 541 GCTTTCATGATTCCTGACCCAGAGCTGCGCTCATAGGCTCATTAATACAGAAATTGCC 600  
 Qy 571 CCTATGAGAGTGGCAAGAAATGCTATGTAAGATGCAAGTGGATTTGTTTCAAGAGAT 630  
 Db 601 CCTATAGAGTGGCTCAAGCTCCAGTAAAGTCCGCTTATTTGTTTGAAGGAATAT 660  
 Qy 631 GCTTCTGCTCGGAGATTAACATATATGCGATCCAGATTTCCTTCCGTATGTTCCG 690  
 Db 661 GGATGACGCTCCGAGATGACATCTATGAGAACCAACTTTTACATTTCCATATGTTTCA 720  
 Qy 691 ACCATTTGCAATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 750  
 Db 721 TCACATCTACATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780  
 Qy 751 CGGTTTGTGACTGATAGAGGTTGACCAACCAATCCGATCATGTTGCTGATGATGATC 810  
 Db 781 CGATATATGAAATTCGATTAAGATGTTCTCCAGTGAATATATGTTGATGAGGGA 840  
 Qy 811 GAAGAATTTAATTAAGATCTCAGATGAGTGAAGTATACGAGAGCGTCTCCCT 870  
 Db 841 GAGATGTAATCTATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
 Qy 871 AAATATTTCACTTACCTTACAGCACTGCAAGAAACCACTTGAAAGATGATGATGATG 930  
 Db 901 AGAATTTTCAATATCTATATAGACATGCAAGAAATCACT-----GATATG 948  
 Qy 931 GGAACCGGTATGTTCCCTGACTATGAGTGGTATGATGATGATGATGATGATGATGATG 990  
 Db 949 GATTCCTCTAGTGAAGAGTAACTATGATGATGATGATGATGATGATGATGATGATG 1008  
 Qy 991 TTGATGCTGCTATTTTGGGAGATTTTGAATATCAATCATGATGATGATGATGATGATG 1050  
 Db 1009 CTGATGCTGATATTTTGGGAGATTTTGAATATCAATCATGATGATGATGATGATG 1068

RESULT 12  
 CD831928 554 bp mRNA linear EST 10-JUL-2003  
 LOCUS CD831928  
 DEFINITION BM40.061P15F011227 BM40 Brassica napus cDNA clone BM40061P15, mRNA sequence.  
 ACCESSION CD831928  
 VERSION CD831928.1 GI:32513868  
 KEYWORDS EST.  
 SOURCE Brassica napus (rape)  
 ORGANISM Brassica napus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE  
 1 (bases 1 to 554)  
 Genoplatane.  
 Genoplatane, a major partnership french program in plant genomics  
 JOURNAL  
 Unpublished (2003)  
 COMMENT  
 Contact: Genoplatane  
 Genoplatane  
 93, rue Henri Rochefort 91025 EVRY CEDEX France  
 Tel: 33 1 69 47 54 00  
 Fax: 33 1 69 47 54 10

FEATURES  
 source  
 1.554  
 /organism="Brassica napus"  
 /mol\_type="mRNA"  
 /cultivar="Jet Neuf"  
 /db\_xref="taxon:3708"  
 /clone="BN40061P15"  
 /tissue\_type="seed"  
 /clone\_id="BN40"  
 location/Qualifiers

## ORIGIN

Query Match 41.2%; Score 455.4; DB 6; Length 554;  
 Best Local Similarity 97.7%; Pred. No. 8.4e-129;  
 Matches 46; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 ATGCGCGTGAAGAAAGCTAGCGAGATGTTTGAAGAGCTTGAAGAGCTTTCACAGA 60  
 DB 81 ATGCGCGTGAAGAAAGCTAGCGAGATGTTTGAAGAGCTTGAAGAGCTTTCACAGA 140  
 QY 61 TGGGAGATGATGAAGCAGACGGGCGTGAAGCTTCAAGATGATGAGATTCGTTCCACT 120  
 DB 141 TGGGAGATGATGAAGCAGACGGGCGTGAAGCTTCAAGATGATGAGATTCGTTCCACT 200  
 QY 121 CCCACTGAGAGAACTTCTGATCTCGGCGAGTTTCTTCAAGAGAGCTTCCGATTCCG 180  
 DB 201 CCCACTGAGAGAACTTCTGATCTCGGCGAGTTTCTTCAAGAGAGCTTCCGATTCCG 260  
 QY 181 ATGCGCGAGCGTGCATGCACTCGAGACGCTGCTTATGCGCTCTCTGAGAAA 240  
 DB 261 ATGCGCGAGCGTGCATGCACTCGAGACGCTGCTTATGCGCTCTCTGAGAAA 320  
 QY 241 GTCTTGAAGTGAAGATTGTATGTGATGATTCATTCAGGACATGAAGGCTTCTCGAG 300  
 DB 321 GTCTTGAAGTGAAGATTGTATGTGATGATTCATTCAGGACATGAAGGCTTCTCGAG 380  
 QY 301 ATCAAGGATCTGCTGATGAGAAAGATTCACACAGATGATCAAGGCTGTTAAAGTAAG 360  
 DB 381 ATCAAGGATCTGCTGATGAGAAAGATTCACACAGATGATTAAGGCTGTTAAAGTAAG 440  
 QY 361 CACAACACGCTGTTCCATGATGCTCTGCGTGTGAACCAAGCTGAAGAAAGTAAGAA 420  
 DB 441 CACAACACGCTGTTCCATGATGCTCTGCGTGTGAACCAAGCTGAAGAAAGTAAGAA 500  
 QY 421 CTCTACGAAAGCTTGAATGATGATTCATGAGTTTCTTGAATGCTTCTTCTGTC 473  
 DB 501 CTCTACGAAAGCTTGAATGATGATTCATGAGTTTCTTGAATGCTTCTTCTGTC 553

RESULT 13  
 BX839466 784 bp mRNA linear EST 11-FEB-2004  
 LOCUS BX839466 Arabidopsis thaliana Adult vegetative tissue Col-0  
 DEFINITION Arabidopsis thaliana cDNA clone GSLTSL342B07 5PRIM, mRNA sequence.  
 ACCESSION BX839466  
 VERSION BX839466.1 GI:42533549  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE  
 1 (bases 1 to 784)  
 Genoplatane.  
 Genoplatane, a major partnership french program in plant genomics  
 JOURNAL  
 Unpublished (2004)  
 COMMENT  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out  
 full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castellil  
 V., Aury J.M., Jailion O., Winkler P., Menard M., Cruaud C.,  
 Schachter V., Weisenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information center for Protein  
 Sequences).  
 http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/EST  
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
 location/Qualifiers

## FEATURES

source

1.784  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /ecotype="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSLTSL342B07"  
 /tissue\_type="Adult vegetative tissue"  
 /clone\_id="Arabidopsis thaliana Adult vegetative tissue  
 Col-0"

## ORIGIN

Query Match 40.8%; Score 450.2; DB 5; Length 784;  
 Best Local Similarity 82.5%; Pred. No. 3.9e-127;  
 Matches 56; Conservative 0; Mismatches 113; Indels 7; Gaps 4;

QY 1 ATGCGCGTGAAGAAAGCTAGCGAGATGTTTGAAGAGCTTGAAGAGCTTTCACAGA 60  
 DB 104 ATGCGAGTGAAGAAAGCTTTCATGTTTCCGAGAGTTGATCGAAGTTTTCACAA 163  
 QY 61 TGGGATGATGATGAAGCAGACGGGCGTGAAGCTTCAAGATGATGAGATTCGTTCCACT 120  
 DB 164 TGGGATGATGATGAAGCAGACGGGCGTGAAGCTTCAAGATGATGAGATTCGTTCCAA 223  
 QY 121 CCCACTGAGAGAACTTCTGATCTCGGCGAGTTTCTTCAAGAGAGCTTCCGATTCCG 180  
 DB 224 CCCACTGAGAGAACTTCTGATCTCGGCGAGTTTCTTCAAGAGAGCTTCCGATTCCG 283  
 QY 181 ATGCGCGTGCATGCACTCGAGACGCTGCTTATGCGCTCTCTGAGAAA 239  
 DB 284 GTGCGTGAAGAGCGATGCACTCGAGACGCTTCTTATGCTCTCTGATTAACCTGC 343  
 QY 240 CGTCTTGAAGTGAAGATTGTATGTGATGATTCATTCAGGACATGAAGGCTTCTCTGA 299  
 DB 344 CGTCTTGAAGTGAAGATTGTATGTGATGATTCATTCAGGACATGAAGGCTTCTCTGA 403  
 QY 300 GATCAAGATACGCTGATGAGAAAGATTCAAC-AGATGATCAAGGCTGTTAAAGTAA 358  
 DB 404 GATTAGAGATTCGGGTGACAGAAAGTTTCACTCAAGATGATTAAGGCTGCAAGTAA 463  
 QY 359 GGCACAAACAGTGTCTCCATGATGCTCTGGGTGTGAACCAAGTGAAGAAAGATGA 418  
 DB 464 GGCATTAACATGATGTTCCATGATGCTTGGGTGTGAATCAAGCTCAAGAAAGTAA 523  
 QY 419 AACTTACGAAAGCTTGAATGATGATTCATGATTTCTTGAATGCTTCTTCTGAT 478  
 DB 524 TTCT---GGAATCTTGAATGATTCATGATTTCTTGAATGCTTCTTCTGATGCTGCGAA 579

QY 479 TAGGATCCGATGCTTA-TCGGGACGATGTTGAGTTCATATACCAACCACACTT 537  
 DB 580 TCGGATCCGGGTCTTATGAGGACGATGTTGATTCATATACCAATCCACGCTT 639  
 QY 538 CACACAGTGGGTATACATACACCAAGATGTCCTCATGAGGTGGCAGGAATGCTGT 597  
 DB 640 CATACAGTGGGTATATACACCAAGATGTCCTCATGAGGTGGCAGGAATGCTGT 699  
 QY 598 GAGATGCAAGTTCGATTTGTTTCAAGATGATGTTGCTCCGAGATTAACATATAT 657  
 DB 700 GAGATGTCGATTCGATTTTTCGAGATGCTGCTCCGAGATTAACATATAT 759  
 QY 658 GCGATCCAAAGTTCACTTTCCGT 682  
 DB 760 GGGATTCGATTTTACCTTTCCGT 784

## RESULT 14

CK208011 1090 bp mRNA linear EST 08-DEC-2003  
 LOCUS FGAS019688 Triticum aestivum FGAS: Library 5 GATE 7 Triticum

DEFINITION aestivum cDNA, mRNA sequence.

ACCESSION CK208011 GI:39570401

VERSION CK208011.1

KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)

ORGANISM Triticum aestivum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

1 (bases 1 to 1090)

Allard, F., Crosby, W. L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,

Genswein, B., Graf, R., Gully, P., Hrycan, L. D., Laroche, A.,

Link, M. G., McCarthy, E. L., Monroy, A., Muzak, I., Nilsson, D.,

Penniket, C., Roach, J. L., and Sarhan, F.

Functional Genomics of Abiotic Stress in Wheat and Canola Crops

Unpublished (2003)

Contact: Wm. L. Crosby

Bioinformatics

University of Saskatchewan, Department of Computer Science

1C101 Engineering Building, 57 Campus Drive, Saskatoon,

Saskatchewan, S7N 5A9, Canada

Tel: 306 966 1769

Fax: 306 966 2033

Email: fgas\_estecs.usask.ca

This sequence is the direct result of the Base calling software

Phred (default parameters). It is the raw base calls. To aid in the

identification of the high quality insert the software Lucy

(default parameters) has been run on this sequence. Lucy identified

the region [25, 811].

Plate: USB012 row: K column: 21.

Location/Qualifiers

1..1090

/organism="Triticum aestivum"

/mol\_type="mRNA"

/db\_xref="taxon:4565"

/note="Vector: pCMV.SPORTS; Crown and developmental stages

of spike formation in wheat cultivar Norstar. 4 mRNA

populations were combined before constructing the library.

The first mRNA population is from 1cm crown sections after

30 days of cold acclimation. The second is from 1cm crown

sections after 11 days of deacclimation (before

deacclimation plants were fully vernalized for 49 days).

The third is from different developmental stages of spike

formation (5 to 50mm) that still have not emerged from the

leaf (dissection required). The last is from different

developmental stages of spike and seed formation after

having emerged from the leaf (visible). First strand

synthesis in this library was done in the presence of

methyated dCTP thereby protecting from internal cleavage

with NotI."

## ORIGIN

Query Match 40.0%; Score 441.6; DB 7; Length 1090;

Best Local Similarity 68.9%; Pred. No. 2e-124;

Matches 639; Conservative 0; Mismatches 276; Indels 13; Gaps 2;

QY 3 GGGGTGAGAGAGGCTTACGAGATGTTTTCGAGACCTTGATGAGAGCTTACAGATG 62  
 DB 100 GACGGGGAGATGAGCGCGGAGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 159  
 QY 63 GGGATGATGAGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 122  
 DB 160 GGGGGGAGAGGAG 219  
 QY 123 CACTGAGAGAAACCTTCTGATCTCGCGGAGATTTCTTCAAGAGAGCTTCGATTCGAT 182  
 DB 220 CACGAGCGGCAACTCTGCTGCTCGCGGAGATTTCTTCAAGAGAGAGAGAGAG 279  
 QY 183 CCGGAGCGGCGGAG 242  
 DB 280 CCGGAGCGGCGGAG 339  
 QY 243 CTGAGAGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 302  
 DB 340 CTTCAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 399  
 QY 303 CAGAGTACTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 362  
 DB 400 GAGCAACGAG 459  
 QY 363 CACAGACGAG 412  
 DB 460 CACAGAGTGGAG 519  
 QY 413 -GAATGAAACTCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 470  
 DB 520 TACAAAGGATTCCTCTCGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGAT 579  
 QY 471 GTCTGTATGAGAGATTCCTGATGATGATGATGATGATGATGATGATGATGAT 530  
 DB 580 GTCAAGATGAGAGATTCCTGATGATGATGATGATGATGATGATGATGATGAT 639  
 QY 531 ACCACTTACACAGAGGAGTATACATACACACCAAGATGCTCTTGAAGAGGAGAG 590  
 DB 640 AGAGCTGCGGATTAAGAGGCTCATTAACCAAGATGCTCTTGAAGAGGAGAG 699  
 QY 591 TGTATGTAAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 650  
 DB 700 TGTATGTAAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 759  
 QY 651 CATATATGCGATCAAGATTCATCTTCCGATGATGATGATGATGATGATGATGAT 710  
 DB 760 CATATATGCGATCAAGATTCATCTTCCGATGATGATGATGATGATGATGATGAT 819  
 QY 711 GTATGATTAAGATCAAGATTCATCTTCCGATGATGATGATGATGATGATGATGAT 770  
 DB 820 GTTGAATGATGAGAGATTCATCTTCCGATGATGATGATGATGATGATGATGAT 879  
 QY 771 GGTGACACCAACATTCGATGATGATGATGATGATGATGATGATGATGATGAT 830  
 DB 880 ACATGACACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 939  
 QY 831 CTCAGATGAAGTGAAGTATACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 890  
 DB 940 TAGTGAAGAGAGTGAAGTATACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 998  
 QY 891 CAGAGTGAAG 918  
 DB 999 CCGTACGAG 1026

RESULT 15  
 CN892416

LOCUS CN892416 669 bp mRNA linear EST 04-JUN-2004  
 DEFINITION 010603AA005454HT (AAXA) Royal Gala 126 DAFB fruit core Malus x  
 domestic cDNA clone AAXA005454, mRNA sequence.  
 ACCESSION CN892416  
 VERSION CN892416.1 GI:48278658  
 KEYWORDS EST.  
 SOURCE Malus x domestica (cultivated apple)  
 ORGANISM Malus x domestica  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids 1; Rosales; Rosaceae; Maloideae; Malus.  
 REFERENCE 1 (bases 1 to 669)  
 Beuning, L., Bowen, J., Crowhurst, R., Gleave, A., Janssen, B.,  
 McArtney, S., Newcomb, R., Ross, G., Snowden, K., Walton, B. and Yauk, Y.  
 HortResearch Apple EST Project  
 TITLE unpublished (2004)  
 JOURNAL  
 COMMENT Contact: Gleave, A.  
 Sequencing Facility  
 The Horticulture and Food Research Institute of New Zealand Ltd  
 120 Mt Albert Rd, Mt Albert, Auckland, New Zealand  
 Tel: 00 64 09 815 4200  
 Fax: 00 64 09 815 4201  
 Email: eest@hortresearch.co.nz  
 FEATURES  
 source  
 1..669  
 Location/Qualifiers  
 /organism="Malus x domestica"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3750"  
 /clone="AAXA005454"  
 /issue\_type="Fruit core"  
 /dev\_stage="126 days after full bloom"  
 /clone\_id="(AAXA) Royal Gala 126 DAFB fruit core"  
 /note="Vector: pBluescript SK(-); library sequenced by  
 Genesis Research & Development"

## ORIGIN

Query Match 37.0%; Score 409; DB 7; Length 669;  
 Best Local Similarity 78.3%; Pred. No. 1.9e-114;  
 Matches 490; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 1 ATGCGCGTGAAGAGGCTGAGATGTTTCCAGAGCTTGATCGAGAGCTTCACAGA 60  
 DB 44 ATGCGCGTTAAGAGGCGAGCGAGTCTTCGAGAGCTTGATCGAGAGCTTCACAGA 103  
 QY 61 TGGGAGATGATGAGAGGAGGCGGCTGAGCTGATGATGAGAGCTTCGATCCACT 120  
 DB 104 TGGGAGATGATGAGAGGAGGCGGCTGAGCTGATGATGAGAGCTTCGATCCACT 163  
 QY 121 CCCACTGAGAGAACTTCTGATCTCGCGCAGTTTCTTCAAGAGAGCTTCCGATTCG 180  
 DB 164 CCGACGAGGCGGATTCATATCTCGCGCAGTTTTCGACAAAGAGCTCCCATTCGG 223  
 QY 181 ATGCGAGGCGTGGATGAGCTCGAGACGCTGCTTATGAGCTCTTGAGAAAGCTGCC 240  
 DB 224 ATGCGAGGCGGCGCAATGAGCTCGAGAGCTCCCTTATGAGCTTGTGAGAAAGCTGCC 283  
 QY 241 GTCTTGAAGGTAGAGATTGATGAGAGTTCAGGAGCAATGAGAGGTTTCCAGAG 300  
 DB 284 GTTTTGAAGGTAGAGATTGATGAGAGTTCAGGAGCAATGAGAGGTTTCCAGAG 343  
 QY 301 ATCAAGGATACCTGCTGATGAGAAAGTTCACAGATGATCAAGGCTGTTAAAGTAAG 360  
 DB 344 ATCAAGGATCTAAAGATGAGAAAGATTTTACCAATGATTAAGGAGATTAAGGTAGA 403  
 QY 361 CACAAACAAGCTGTTCCCATGATGCTCTGGGTGTGAACGAGTGAAGAAAGAGTAAA 420  
 DB 404 CACAACAATGTTGCTCCGATGATGCTTGGGTGTTCAGCAGTTGAAGAAAGAGAGTT 463  
 QY 421 CTCTAGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
 DB 464 CTCTAGAGAGATCTTGAATGAGAAATTCACCAATTTTGAATGCTTCTACATGTCAGAAAT 523  
 QY 481 GGGATCGTATGCTTATCGGCGAGCATGTTGATGATGATGATGATGATGATGATGAT 540

DB 524 GGGATTCGATGCTTCAATGGCCAGCATGTTGAGTTGACATCCCATCCCTCCAT 583  
 QY 541 ACAGTGGTTACATACACACCAAGATGCTTCCATGAGAGGTGCAAGGAATGCTAGTAA 600  
 DB 584 TGTGTGGTTATATAGATACAAAGATGCTCCCATGAGGTAGACAGAAATGCCACTGAG 643  
 QY 601 GATCAAGGTGATGTTGTTTCAAGAGA 626  
 DB 644 GAGCCCGTGCATGTCCTGCGGTGA 669

Search completed: April 12, 2005, 08:04:50  
 Job time : 4258 secs

***This Page Blank (uspto)***